Improved Fermentative Carotenoid Production

Background of the Invention

Over 600 different carotenoids have been described from carotenogenic organisms found among bacteria, yeast, fungi and plants. Currently only two of them, β-carotene and astaxanthin are commercially produced in microorganisms and used in the food and feed industry. β-carotene is obtained from algae and astaxanthin is produced in Pfaffia strains which have been generated by classical mutation. However, fermentation in Pfaffia has the disadvantage of long fermentation cycles and recovery from algae is cumbersome. Therefore it is desirable to develop production systems which have better industrial applicability, e.g. can be manipulated for increased titers and/or reduced fermentation times.

Two such systems using the biosynthetic genes form Erwinia herbicola and Erwinia uredovora have already been described in WO 91/13078 and EP 393 690, respectively. Furthermore, three β-carotene ketolase genes (β-carotene β-4-oxygenase) of the marine bacteria *Agrobacterium aurantiacum* and *Alcaligenes* strain PC-1 (crtW) [Misawa, 1995, Biochem. Biophys. Res. Com. 209, 867-876][Misawa, 1995, J. Bacteriology 177, 6575-6584] and from the green algae *Haematococcus pluvialis* (bkt) [Lotan, 1995, FEBS Letters 364, 125-128][Kajiwara, 1995, Plant Mol. Biol. 29, 343-352] have been cloned. *E. coli* carrying either the carotenogenic genes (crtE, crtB, crtY and crtI) of *E. herbicola* [Hundle, 1994, MGG 245, 406-416] or of *E. uredovora* and complemented with the crtW gene of *A. aurantiacum* [Misawa, 1995] or the bkt gene of *H. pluvialis* [Lotan, 1995][Kajiwara, 1995] resulted in the accumulation of canthaxanthin (β,β-carotene-4,4'-dione), originating from the conversion of β-carotene, via the intermediate echinenone (β,β-carotene-4-one).

Introduction of the above mentioned genes (crtW or bkt) into *E. coli* cells harbouring besides the carotenoid biosynthesis genes mentioned above also the crtZ gene of *E. uredovora* [Kajiwara, 1995][Misawa, 1995], resulted in both cases in the accumulation of astaxanthin (3,3'-dihydroxy-β,β-carotene-4,4'-dione). The

results obtained with the bkt gene, are in contrast to the observation made by others [Lotan, 1995], who using the same experimental set-up, but introducing the *H. pluvialis* bkt gene in a zeaxanthin (β,β-carotene-3,3'-diol) synthesising *E. coli* host harbouring the carotenoid biosynthesis genes of *E. herbicola*, a close relative of the above mentioned *E. uredovora* strain, did not observe astaxanthin production.

Summary of the Invention

Novel proteins of microorganism E-396 (FERM BP-4283) and the DNA sequences which encode these proteins have been discovered which provide an improved biosynthetic pathway from farnesyl pyrophosphate and isopentyl pyrophosphate to various carotenoids, especially zeaxanthin, astaxanthin, adonixanthin and canthaxanthin.

Brief Description of the Figures

Fig. 1: The biosynthesis pathway for the formation or carotenoids of *Flavobacterium* sp. R1534 is illustrated explaining the enzymatic activities which are encoded by DNA sequences of the present invention.

Fig. 2: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with the restriction enzymes shown on top of each lane and hybridized with Probe 46F. The arrow indicated the isolated 2.4 kb XhoI/PstI fragment.

Fig. 3: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with ClaI or double digested with ClaI and HindIII. Blots shown in Panel A and B were hybridized to probe A or probe B, respectively (see examples). Both ClaI/HindIII fragments of 1.8 kb and 9.2 kb are indicated.

Fig. 4: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with the restriction enzymes shown on top of each lane and hybridized to probe C. The isolated 2.8 kb SalI/HindIII fragment is shown by the arrow.

Fig. 5: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested

with the restriction enzymes shown on top of each lane and hybridized to probe D. The isolated BclI/SphI fragment of approx. 3 kb is shown by the arrow.

Fig. 6: Physical map of the organization of the carotenoid biosynthesis cluster in Flavobacterium sp. R1534, deduced from the genomic clones obtained. The location of the probes used for the screening are shown as bars on the respective clones.

Fig. 7: Nucleotide sequence of the *Flavobacterium* sp. R1534 carotenoid biosynthesis cluster and its flanking regions (SEQ ID NO: 1). The nucleotide sequence is numbered from the first nucleotide shown (see BamHI site of Fig. 6). The deduced amino acid sequence of the ORF's (orf-5, orf-1, crtE, crtB, crtI, crtY, crtZ and orf-16) are shown with the single-letter amino acid code. Arrow (-->) indicate the direction of the transcription; asterisks, stop codons.

Fig. 8: Protein sequence of the GGPP synthase (crtE) of *Flavobacterium* sp. R1534 (SEQ ID NO: 2) with a MW of 31331 Da.

Fig. 9: Protein sequence of the prephytoene synthetase (crtB) of *Flavobacterium* sp. R1534 (SEQ ID NO: 3) with a MW of 32615 Da.

Fig. 10: Protein sequence of the phytoene desaturase (crtI) of *Flavobacterium* sp. R1534 (SEQ ID NO: 4) with a MW of 54411 Da.

Fig. 11: Protein sequence of the lycopene cyclase (crtY) of *Flavobacterium* sp. R1534 (SEQ ID NO: 5) with a MW of 42368 Da.

Fig. 12: Protein sequence of the β-carotene hydroxylase (crtZ) of *Flavobacterium* sp. R1534 (SEQ ID NO: 6) with a MW of 19282 Da.

Fig. 13: Recombinant plasmids containing deletions of the *Flavobacterium* sp. R1534 carotenoid biosynthesis gene cluster.

Fig. 14: Primers used for PCR reactions (SEQ ID NOs: 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18). The underlined sequence is the recognition site of the indicated restriction enzyme. Small caps indicate nucleotides introduced by mutagenesis. Boxes show the artificial RBS which is recognized in B. subtilis. Small caps in bold show the location of the

original adenine creating the translation start site (ATG) of the following gene (see original operon). All the ATG's of the original Flavobacter carotenoid biosynthetic genes had to be destroyed to not interfere with the rebuild transcription start site. Arrows indicate start and ends of the indicated Flavobacterium R1534 WT carotenoid genes.

Fig. 15: Linkers used for the different constructions (SEQ ID NOs: 19, 20, 21, 22, 23, 24, 25, and 26). The underlined sequence is the recognition site of the indicated restriction enzyme. Small caps indicate nucleotides introduced by synthetic primers. Boxes show the artificial RBS which is recognized in B. subtilis. Arrow indicate start and ends of the indicated Flavobacterium carotenoid genes.

Fig. 16: Costruction of plasmids pBIIKS(+)-clone59-2, pLyco and pZea4.

Fig. 17: Construction of plasmid p602CAR.

Fig. 18: Construction of plasmids pBIIKS(+)-CARVEG-E and p602 CARVEG-E.

Fig. 19: Construction of plasmids pHP13-2CARZYIB-EINV and pHP13-2PN25ZYIB-EINV.

Fig. 20: Construction of plasmid pXI12-ZYIB-EINVMUTRBS2C.

Fig. 21: Northern blot analysis of B. subtilis strain BS1012::ZYIB-EINV4. Panel A: Schematic representation of a reciprocal integration of plasmid pXI12-ZYIB-EINV4 into the levan-sucrase gene of B- subtilis. Panel B: Northern blot obtained with probe A (PCR fragment which was obtained with CAR 51 and CAR 76 and hybridizes to the 3' end of crtZ and the 5' end or crtY). Panel C: Northern blot obtained with probe B (BamHI-XhoI fragment isolated from plasmid pBIIKS(+)-crtE/2 and hybridizing to the 5' part of the crtE gene).

Fig. 22: Schematic representation of the integration sites of three transformed Bacillus subtilis strains: BS1012::SFCO, BS1012::SFCOCAT1 and BA1012::SFCONEO1. Amplification of the synthetic Flavobacterium carotenoid operon (SFCO) can only be obtained in those strains having amplifiable structures. Probe A was used to determine the copy

number of the integrated SFCO. Erythromycine resistance gene (ermAM), chloramphenicol resistance gene (cat), neomycine resistance gene (neo), terminator of the cryT gene of B. subtilis (cryT), levan-sucrase gene (sac-B 5' and sac-B 3'), plasmid sequences of pXI12 (pXI12), promoter originating from site I of the veg promoter complex (Pvegl).

Fig. 23: Construction of plasmids pXI12-ZYIB-EINV4MUTRBS2CNEO and pXI12-ZYIB-EINV4MUTRBS2CCAT.

Fig. 24: Complete nucleotide sequence of plasmid pZea4 (SEQ ID NO: 27).

Fig. 25: Synthetic crtW gene of Alcaligenes PC-1 (SEQ ID NO: 28). The translated protein sequence (SEQ ID NO: 29) is shown above the double stranded DNA sequence. The twelve oligonucleotides (crtW1-crtW12) used for the PCR synthesis are underlined.

Fig. 26: Construction of plasmid pBIIKS-crtEBIYZW. The HindIII-Pm1I fragment of pALTER-Ex2-crtW, carrying the synthetic crtW gene, was cloned into the HindIII and MluI (blunt) sites. PvegI and Ptac are the promoters used for the transcription of the two opera. The ColE1 replication origin of this plasmid is compatible with the p15A origin present in the pALTER-Ex2 constructs.

Fig. 27: Relevant inserts of all plasmids constructed in Example 7. Disrupted genes are shown by //. Restriction sites: S=SacI, X=Xbal, H=HindIII, N=NsiI, Hp=HpaI, Nd=NdeI.

Fig. 28: Reaction products (carotenoids) obtained from β-carotene by the process of the present invention.

Fig. 29: Isolation of the crt cluster of the strain E- 396. Genomic DNA of E-396 was digested overnight with different combinations of restrictions enzymes and separated by agarose gel electrophoresis before transferring the resulting fragments by Southern blotting onto a nitrocellulose membrane. The blot was hybridised with a ^{32}P labelled 334 bp fragment obtained by digesting the aforementioned PCR fragment JAPclone8 with *Bss*HII and *Mlu*I. An approx. 9,4kb *Eco*RI/*Bam*HI fragment hybridizing to the probe was identified as the most appropiate for cloning since it is long enough to potentially carry the complete *crt* cluster. The fragment was isolated

and cloned into the *EcoR*I and *BamH*I sites of pBluescriptIIKS resulting in plasmid pJAPCL544.

Fig. 30 Shows the sequence obtained containing the $crtW_{E396}$ (from nucleotide 40 to 768) and $crtZ_{E396}$ (from nucleotide 765 to 1253) genes of the bacterium E-396 (SEQ ID NO: 30).

Fig. 31: The nucleotide sequence of the $crtW_{E396}$ gene (SEQ ID NO: 31)

Fig. 32: The amino acid sequence encoded by the $crtW_{E396}$ (SEQ ID NO: 32) gene shown in Fig. 31

Fig 33: The nucleotide sequence of the $crtZ_{E396}$ (SEQ ID NO: 33) gene

Fig. 34: The amino acid sequence (SEQ ID NO: 34) encoded by the $crtZ_{E396}$ gene shown in Fig. 33

Fig. 35: Diagram of plasmid pUC18-E396crtWZPCR

Fig. 36: Construction of plasmid pBIIKS-crtEBIY[E396WZ]

Fig. 37: Construction of plasmid pBIIKS-crtEBIY[E396W]DZ which has a truncated non-functional crtZ gene

Fig. 38: 463 bp *Pst*I-*BamH*I fragment (SEQ ID NO: 35) originating from the 3′ end of the insert of pJAPCL544 (Fig. 29) highlighted a ~1300 bp-long *Pst*I-*Pst*I fragment. This fragment was isolated and cloned into the *Pst*I site of pBSIIKS(+) resulting in plasmid pBSIIKS-#1296. The sequence of the insert is shown (small cap letters refer to new sequence obtained. Capital letters show the sequence also present in the 3′ of the insert of plasmid pJAPCL544).

Fig. 39: The DNA sequence of the complete $crtE_{E396}$ gene (SEQ ID NO: 36)

Fig. 40: The amino acid sequence encoded by the $crtE_{E396}$ gene (SEQ ID NO: 37) shown in Fig. 39 (SEQ ID NO: 36)

Fig. 41: Construction of plasmid carrying the complete *crt* cluster of E-396 (pE396CARcrtW-E)

Fig. 42: Construction of plasmid pRSF1010-Amp^r

Fig. 43: Construction of plasmids RSF1010-Amp^r-crt1 and RSF1010-Ampr-crt2

Detailed Description of the Invention

Novel proteins of microorganism E-396 (FERM BP-4283) and the DNA sequences which encode these proteins have been discovered which provide an improved biosynthetic pathway from farnesyl pyrophosphate and isopentyl pyrophosphate to various carotenoids, especially zeaxanthin, astaxanthin, adonixanthin and canthaxanthin.

One aspect of the invention is a polynucleotide comprising a DNA sequence which encodes the GGPP synthase ($crtE_{E396}$) (SEQ ID NO: 37) of microorganism E-396, said polynucleotide being substantially free of other polynucleotides of microorganism E-396. Also encompassed by this aspect of the present invention is a polynucleotide comprising a DNA sequence which is substantially homologous to said DNA sequence. Said GGPP synthase catalyzes the condensation of farnesyl pyrophosphate and isopentyl pyrophosphate to obtain geranylgeranyl pyrophosphate, a carotenoid precursor. The preferred GGPP synthase has the amino acid sequence of Figure 40 (SEQ ID NO: 37), and the preferred DNA sequence encodes said amino acid sequence. The especially preferred DNA sequence is shown in Figure 39 (SEQ ID NO: 36).

This aspect of the present invention also includes a vector comprising the aforesaid polynucleotide, preferably in the form of an expression vector. Furthermore this aspect of the present invention also includes a recombinant cell comprising a host cell which is transformed by the aforesaid polynucleotide or vector which contains such a polynucleotide. Preferably said host cell is a prokaryotic cell and more preferably said host cell is *E. coli* or a *Bacillus* strain. However, said host cell may also be a eukaryotic cell, preferably a yeast cell or a fungal cell.

Finally this aspect of the present invention also comprises a process for the preparation of geranylgeranyl pyrophosphate by culturing said recombinant cell of the invention containing farnesyl pyrophosphate and isopentyl pyrophosphate in a culture medium under suitable culture conditions whereby said GGPP synthase is expressed by said cell and catalyzes the condensation of farnesyl pyrophosphate and isopentyl pyrophosphate to geranylgeranyl pyrophosphate, and isolating the geranylgeranyl pyrophosphate from such cells or the culture medium.

Another aspect of the present invention is a polynucleotide comprising a DNA sequence which encodes said ß-carotene hydroxylase of microorganism E-396 ($crtZ_{E396}$) (SEQ ID NO: 34), said polynucleotide being substantially free of other polynucleotides of microorganism E-396. Also encompassed by this aspect of the present invention is a polynucleotide comprising a DNA sequence which is substantially homologous to said DNA sequence. Said ß-carotene hydroxylase catalyzes the hydroxylation of ß-carotene to produce the xanthophyll, zeaxanthin. The preferred ß-carotene hydroxylase has the amino acid sequence of Figure 34 (SEQ ID NO: 34), and the preferred DNA sequence is one which encodes said amino acid sequence. The especially preferred DNA sequence is a DNA sequence comprising the sequence shown in Figure 33 (SEQ ID NO: 33).

This aspect of the present invention also includes a vector comprising the aforesaid polynucleotide, preferably in the form of an expression vector. Furthermore this aspect of the present invention also includes a recombinant cell comprising a host cell which is transformed by the aforesaid polynucleotide or vector which contains such a DNA sequence. Preferably said host cell is a prokaryotic cell and more preferably said host cell is *E. coli* or a *Bacillus* strain. However, said host cell may also be a eukaryotic cell, preferably a yeast cell or a fungal cell.

Finally this aspect of the present invention also comprises a process for the preparation of zeaxanthin by culturing said recombinant cell of the invention containing ß-carotene in a culture medium under suitable culture conditions whereby said ß-carotene hydroxylase is expressed by said cell and catalyzes the hydroxylation of ß-carotene to produce the xanthophyll, zeaxanthin, and isolating the zeaxanthin from such cells or the culture medium.

Another aspect of the present invention is a polynucleotide comprising a DNA sequence which encodes said ß-carotene hydroxylase of microorganism E-396 ($crtW_{E396}$) (SEQ ID NO: 32), said polynucleotide being substantially free of other polynucleotides of microorganism E-396. Also encompassed by this aspect of the present invention is a polynucleotide comprising a DNA sequence which is substantially homologous to said DNA sequence. Said ß-carotene ß4-oxygenase catalyzes the hydroxylation of ß-carotene to produce the echinenone, and, with the further catalysis of echinenone by the enzyme encoded by $crtW_{E396}$, to canthaxanthin. The preferred ß-carotene ß4-oxygenase has the amino acid sequence of Figure 32 (SEQ ID NO: 32), and the preferred DNA sequence is one

which encodes said amino acid sequence. The especially preferred DNA sequence is a DNA sequence comprising the sequence shown in Figure 31 (SEQ ID NO: 31).

This aspect of the present invention also includes a vector comprising the aforesaid polynucleotide, preferably in the form of an expression vector. Furthermore this aspect of the present invention also includes a recombinant cell comprising a host cell which is transformed by the aforesaid polynucleotide or vector which contains such a DNA sequence. Preferably said host cell is a prokaryotic cell and more preferably said host cell is *E. coli* or a *Bacillus* strain. However, said host cell may also be a eukaryotic cell, preferably a yeast cell or a fungal cell.

Finally this aspect of the present invention also comprises a process for the preparation of canthaxanthin by culturing said recombinant cell of the invention containing ß-carotene in a culture medium under suitable culture conditions whereby said ß-carotene ß4-oxygenase is expressed by said cell and catalyzes the conversion of ß-carotene to produce echinenone and through further catalysis to produce canthaxanthin, and isolating the canthaxanthin from such cells or the culture medium.

It is contemplated, and in fact preferred, that the aforementioned DNA sequences, $crtE_{E396}$, $crtW_{E396}$ and $crtZ_{E396}$, which terms refer to the above-described genes of microorganism E-396 encompassed by the invention herein described, are incorporated, especially $crtW_{E396}$ and $crtZ_{E396}$, with selected DNA sequences from Flavobacterium sp. R1534 into a polynucleotide of the invention whereby two or more of said DNA sequences which encode enzymes catalyzing contiguious steps in the process shown in Figures 1 and 28 are contained in said polynucleotide, said polynucleotide being substantially free of other polynucleotides of microorganism E-396 and Flavobacterium sp. R1534, to obtain advantageous production of the carotenoids canthaxanthin, zeaxanthin, astaxanthin and adonixanthin.

Thus, one embodiment of the present invention is a process for the preparation of zeaxanthin which process comprises culturing a recombinant cell containing farnesyl pyrophosphate and isopentyl pyrophosphate under culture conditions sufficient for the expression of enzymes which catalyze the conversion of the farnesyl pyrophosphate and isopentyl pyrophosphate to zeaxanthin, said recombinant cell comprising a host cell transformed by an expression vector

comprising a regulatory sequence and a polynucleotide containing DNA sequences which encode said enzymes, as follows:

a) a DNA sequence which encodes the GGPP synthase of Flavobacterium sp. R1534 (crtE) (SEQ ID NO: 2) or a DNA sequence which is substantially homologous,

b) a DNA sequence which encodes the prephytoene synthase of Flavobacterium sp. R1534 (crtB) (SEQ ID NO: 3) or a DNA sequence which is substantially homologous,

c) a DNA sequence which encodes the phytoene desaturase of Flavobacterium sp. R1534 (crtI) (SEQ ID NO: 4) or a DNA sequence which is substantially homologous,

d) a DNA sequence which encodes the lycopene cyclase of Flavobacterium sp. R1534 (crtY) (SEQ ID NO: 5) or a DNA sequence which is substantially homologous,

e) a DNA sequence which encodes the ß-carotene hydroxylase of microorganism E-396 ($crtZ_{E396}$) (SEQ ID NO: 34) or a DNA sequence which is substantially homologous;

and isolating the zeaxanthin from such cells or the culture medium.

The above-described polynucleotide encodes enzymes which catalyze the conversion of farnesyl pyrophosphate and isopentyl pyrophosphate to zeaxanthin. It is preferred that this embodiment of the invention utilize a polynucleotide containing crtE, crtB, crtI, crtY, and $crtZ_{E396}$.

It is especially preferred that for this embodiment of the invention:

a) the GGPP synthase has the amino acid sequence of Figure 8 (SEQ ID NO: 2),

b) the prephytoene synthase has the amino acid sequence of Figure 9 (SEQ ID NO: 3),

c) the phytoene desaturase has the amino acid sequence of Figure 10 (SEQ ID NO: 4),

d) the lycopene cyclase has the amino acid sequence of Figure 11 (SEQ ID NO: 5), and

e) the ß-carotene hydroxylase has the amino acid sequence of Figure 34.

It is most preferred that for this embodiment of the invention:

a) the DNA sequence encoding the GGPP synthase comprises bases 2521-3408 of Figure 7 (SEQ ID NO: 1),

b) the DNA sequence encoding the prephytoene synthase comprises bases 4316-3405 of Figure 7 (SEQ ID NO: 1),

c) the DNA sequence encoding the phytoene desaturase comprises bases 4313-5797 of Figure 7 (SEQ ID NO: 1),

d) the DNA sequence encoding the lycopene cyclase comprises bases 5794-6942 of Figure 7 (SEQ ID NO: 1), and

e) the DNA sequence encoding the ß-carotene hydroxylase comprises the sequence of Figure 33 (SEQ ID NO: 33).

A second embodiment of the invention is a process for the preparation of canthaxanthin which process comprises culturing a recombinant cell containing farnesyl pyrophosphate and isopentyl pyrophosphate under culture conditions sufficient for the expression of enzymes which catalyze the conversion of the farnesyl pyrophosphate and isopentyl pyrophosphate to canthaxanthin, said recombinant cell comprising a host cell transformed by an expression vector comprising a regulatory sequence and a polynucleotide containing DNA sequences which encode said enzymes, as follows:

a) a DNA sequence which encodes the GGPP synthase of Flavobacterium sp. R1534 (crtE) (SEQ ID NO: 2) or a DNA sequence which is substantially homologous,

b) a DNA sequence which encodes the prephytoene synthase of Flavobacterium sp. R1534 (crtB) (SEQ ID NO: 3) or a DNA sequence which is substantially homologous,

c) a DNA sequence which encodes the phytoene desaturase of Flavobacterium sp. R1534 (crtI) (SEQ ID NO: 4) or a DNA sequence which is substantially homologous,

d) a DNA sequence which encodes the lycopene cyclase of Flavobacterium sp. R1534 (crtY) (SEQ ID NO: 5) or a DNA sequence which is substantially homologous, and

e) a DNA sequence which encodes the ß-carotene ß4-oxygenase of microorganism E-396 ($crtW_{E396}$) (SEQ ID NO: 32) or a DNA sequence which is substantially homologous;

and isolating the canthaxanthin from such cells or the culture medium.

The above-described polynucleotide encodes enzymes which catalyze the conversion of farnesyl pyrophosphate and isopentyl pyrophosphate to canthaxanthin. It is preferred that this embodiment of the invention utilize a polynucleotide containing crtE, crtB, crtI, crtY, and $crtW_{E396}$.

It is especially preferred that for this embodiment of the invention:

a) the GGPP synthase has the amino acid sequence of Figure 8 (SEQ ID NO: 2),

b) the prephytoene synthase has the amino acid sequence of Figure 9 (SEQ ID NO: 3),

c) the phytoene desaturase has the amino acid sequence of Figure 10 (SEQ ID NO: 4),

d) the lycopene cyclase has the amino acid sequence of Figure 11 (SEQ ID NO: 5), and

e) the ß-carotene ß4-oxygenase has the amino acid sequence of Figure 32 (SEQ ID NO: 32).

For this embodiment of the invention, it is most preferred that:

a) the DNA sequence encoding the GGPP synthase comprises bases 2521-3408 of Figure 7 (SEQ ID NO: 1),

b) the DNA sequence encoding the prephytoene synthase comprises bases 4316-3405 of Figure 7 (SEQ ID NO: 1),

c) the DNA sequence encoding the phytoene desaturase comprises bases 4313-5797 of Figure 7 (SEQ ID NO: 1),

d) the DNA sequence encoding the lycopene cyclase comprises bases 5794-6942 of Figure 7 (SEQ ID NO: 1), and

e) the DNA sequence encoding the ß-carotene ß4-oxygenase comprises the sequence of Figure 31.

A third embodiment of the invention is a process for the preparation of

astaxanthin and adonixanthin wherein said process comprises culturing a recombinant cell containing farnesyl pyrophosphate and isopentyl pyrophosphate under culture conditions sufficient for the expression of enzymes which catalyze the conversion of the farnesyl pyrophosphate and isopentyl pyrophosphate to astaxanthin and adonixanthin, said recombinant cell comprising a host cell transformed by an expression vector comprising a regulatory sequence and a polynucleotide containing DNA sequences which encode said enzymes, as follows:

a) a DNA sequence which encodes the GGPP synthase of Flavobacterium sp. R1534 (crtE) (SEQ ID NO: 2) or a DNA sequence which is substantially homologous,

b) a DNA sequence which encodes the prephytoene synthase of Flavobacterium sp. R1534 (crtB) (SEQ ID NO: 3) or a DNA sequence which is substantially homologous,

c) a DNA sequence which encodes the phytoene desaturase of Flavobacterium sp. R1534 (crtI) (SEQ ID NO: 4) or a DNA sequence which is substantially homologous,

d) a DNA sequence which encodes the lycopene cyclase of Flavobacterium sp. R1534 (crtY) (SEQ ID NO: 5) or a DNA sequence which is substantially homologous,

e) a DNA sequence which encodes the ß-carotene b4-oxygenase of Flavobacterium sp. R1534 (crtW) or a DNA sequence which is substantially homologous, and

f) a DNA sequence which encodes the ß-carotene hydroxylase of microorganism E-396 ($crtZ_{E396}$) or a DNA sequence which is substantially homologous;

and isolating the astaxanthin and adonixanthin from such cells or the culture medium.

The above-described polynucleotide encodes enzymes which catalyze the conversion of farnesyl pyrophosphate and isopentyl pyrophosphate to astaxanthin and adonixanthin. It is preferred that this embodiment of the invention utilize a polynucleotide containing crtE, crtB, crtI, crtY, crtW, and $crtZ_{E396}$ (SEQ ID NO: 34).

It is especially preferred that for this embodiment of the invention:

a) the GGPP synthase has the amino acid sequence of Figure 8 (SEQ ID NO: 2),

b) the prephytoene synthase has the amino acid sequence of Figure 9 (SEQ ID NO: 3),

c) the phytoene desaturase has the amino acid sequence of Figure 10 (SEQ ID NO: 4),

d) the lycopene cyclase has the amino acid sequence of Figure 11 (SEQ ID NO: 5),

e) the ß-carotene ß4-oxygenase has the amino acid sequence of Figure 25 (SEQ ID NO: 29), and

f) the ß-carotene hydroxylase has the amino acid sequence of Figure 34 (SEQ ID NO: 34).

It is most preferred that for this embodiment of the invention:

a) the DNA sequence encoding the GGPP synthase comprises bases 2521-3408 of Figure 7 (SEQ ID NO: 1),

b) the DNA sequence encoding the prephytoene synthase comprises bases 4316-3405 of Figure 7 (SEQ ID NO: 1),

c) the DNA sequence encoding the phytoene desaturase comprises bases 4313-5797 of Figure 7 (SEQ ID NO: 1),

d) the DNA sequence encoding the lycopene cyclase comprises bases 5794-6942 of Figure 7 (SEQ ID NO: 1),

e) the DNA sequence encoding the ß-carotene ß4-oxygenase comprises the sequence of Figure 25 (SEQ ID NO: 28), and

f) the DNA sequence encoding the ß-carotene hydroxylase comprises the sequence of Figure 33 (SEQ ID NO: 33).

A fourth embodiment of the invention is a process for the preparation of astaxanthin and adonixanthin wherein said process comprises culturing a recombinant cell containing farnesyl pyrophosphate and isopentyl pyrophosphate under culture conditions sufficient for the expression of enzymes which catalyze the conversion of the farnesyl pyrophosphate and isopentyl pyrophosphate to astaxanthin and adonixanthin, said recombinant cell

comprising a host cell transformed by an expression vector comprising a regulatory sequence and a polynucleotide containing DNA sequences which encode said enzymes, as follows:

a) a DNA sequence which encodes the GGPP synthase of Flavobacterium sp. R1534 (crtE) (SEQ ID NO: 2) or a DNA sequence which is substantially homologous,

b) a DNA sequence which encodes the prephytoene synthase of Flavobacterium sp. R1534 (crtB) (SEQ ID NO: 3) or a DNA sequence which is substantially homologous,

c) a DNA sequence which encodes the phytoene desaturase of Flavobacterium sp. R1534 (crtI) (SEQ ID NO: 4) or a DNA sequence which is substantially homologous,

d) a DNA sequence which encodes the lycopene cyclase of Flavobacterium sp. R1534 (crtY) (SEQ ID NO: 5) or a DNA sequence which is substantially homologous,

e) a DNA sequence which encodes the ß-carotene ß4-oxygenase of microorganism E-396 ($crtW_{E396}$) (SEQ ID NO: 32) or a DNA sequence which is substantially homologous, and

f) a DNA sequence which encodes the ß-carotene hydroxylase of microorganism E-396 ($crtZ_{E396}$) (SEQ ID NO: 34) or a DNA sequence which is substantially homologous;

and isolating the astaxanthin and adonixanthin from such cells or the culture medium.

The above-described polynucleotide encodes enzymes which catalyze the conversion of farnesyl pyrophosphate and isopentyl pyrophosphate to astaxanthin and adonixanthin. It is preferred that this embodiment of the invention utilize a polynucleotide containing crtE, crtB, crtI, crtY, $crtW_{E396}$, and $crtZ_{E396}$.

It is especially preferred that for this embodiment of the invention:

a) the GGPP synthase has the amino acid sequence of Figure 8 (SEQ ID NO: 2),

b) the prephytoene synthase has the amino acid sequence of Figure 9 (SEQ ID NO: 3),

c) the phytoene desaturase has the amino acid sequence of Figure 10 (SEQ ID NO: 4),

d) the lycopene cyclase has the amino acid sequence of Figure 11 (SEQ ID NO: 5),

e) the ß-carotene ß4-oxygenase has the amino acid sequence of Figure 32 (SEQ ID NO: 32), and

f) the ß-carotene hydroxylase has the amino acid sequence of Figure 34 (SEQ ID NO: 34).

It is most preferred that for this embodiment of the invention:

a) the DNA sequence encoding the GGPP synthase comprises bases 2521-3408 of Figure 7 (SEQ ID NO: 1),

b) the DNA sequence encoding the prephytoene synthase comprises bases 4316-3405 of Figure 7 (SEQ ID NO: 1),

c) the DNA sequence encoding the phytoene desaturase comprises bases 4313-5797 of Figure 7 (SEQ ID NO: 1),

d) the DNA sequence encoding the lycopene cyclase comprises bases 5794-6942 of Figure 7 (SEQ ID NO: 1),

e) the DNA sequence encoding the ß-carotene ß4-oxygenase comprises the sequence of Figure 31 (SEQ ID NO: 31), and

f) the DNA sequence encoding the ß-carotene hydroxylase comprises the sequence of Figure 33 (SEQ ID NO: 33).

A fifth embodiement of the present invention is a process for the preparation of adonixanthin wherein said process comprises culturing a recombinant cell containing farnesyl pyrophosphate and isopentyl pyrophosphate under culture conditions sufficient for the expression of enzymes which catalyze the conversion of the farnesyl pyrophosphate and isopentyl pyrophosphate to adonixanthin, said recombinant cell comprising a host cell transformed by an expression vector comprising a regulatory sequence and a polynucleotide containing DNA sequences which encode said enzymes, as follows:

a) a DNA sequence which encodes the GGPP synthase of microorganism E-396 ($crtE_{E396}$) (SEQ ID NO: 37) or a DNA sequence which is substantially homologous,

b) a DNA sequence which encodes the prephytoene synthase of microorganism E-396 ($crtB_{E396}$) or a DNA sequence which is substantially homologous,

c) a DNA sequence which encodes the phytoene desaturase of microorganism E-396 ($crtI_{E396}$) or a DNA sequence which is substantially homologous,

d) a DNA sequence which encodes the lycopene cyclase of microorganism E-396 ($crtY_{E396}$) or a DNA sequence which is substantially homologous,

e) a DNA sequence which encodes the b-carotene b4-oxygenase of microorganism E-396 ($crtW_{E396}$) (SEQ ID NO: 32) or a DNA sequence which is substantially homologous, and

f) a DNA sequence which encodes the β-carotene hydroxylase of microorganism E-396 ($crtZ_{E396}$) (SEQ ID NO: 33) or a DNA sequence which is substantially homologous,

said host cell being substantially free of other polynucleotides of microorganism E-396;

and isolating the adonixanthin from such cells or the culture medium.

The above-described polynucleotide encodes enzymes which catalyze the conversion of farnesyl pyrophosphate and isopentyl pyrophosphate to adonixanthin. It is preferred that this embodiment of the invention utilize a polynucleotide containing $crtE_{E396}$, $crtB_{E396}$, $crtI_{E396}$, $crtY_{E396}$, $crtW_{E396}$, and $crtZ_{E396}$. It has been found that the use of the above-described process of the invention results in a preferential production of adonixanthin in relation to astaxanthin and other carotenoids. The preferred polynucleotide is plasmid pE396CARcrtW-E whose construction is described in Example 9 herein.

The present invention also comprises the polynucleotides described above for the various embodiments of the invention and a vector comprising such a polynucleotide, preferably in the form of an expression vector. Furthermore the present invention also comprises a recombinant cell wherein said cell is a host cell

which is transformed by a polynucleotide of the invention or vector which contains such a polynucleotide. Host cells useful for the expression of heterologous genes normally contain farnesyl pyrophosphate and isopentyl pyrophosphate, which are used for other purposes within the cell. Preferably said host cell is a prokaryotic cell and more preferably said host cell is an *E. coli* or a *Bacillus* strain. However, said host cell may also be a eukaryotic cell, preferably a yeast cell or a fungal cell.

Finally the present invention also comprises a process for the preparation of a desired carotenoid by culturing a recombinant cell of the invention containing a starting material in a culture medium under suitable culture conditions and isolating the desired carotenoid from such cells or the culture medium wherein the cell utilizes the polynucleotide of the invention which contains said DNA sequences to express the enzymes which catalyze the reactions necessary to produce the desired carotenoid from the starting material. Where an enzyme catalyzes two sequential steps and it is preferred to produce the product of the second step (such as producing astaxanthin preferentially to adonixanthin (see Fig. 28)), a higher copy number of the DNA sequence encoding the enzyme may be used to further production of the product of the second of the two steps in comparison to the first product. The present invention further comprises a process for the preparation of a food or feed composition which process comprises mixing a nutritionally effective amount of the carotenoid isolated from the aforementioned recombinant cells or culture medium with said food or feed.

In this context it should be mentioned that the expression "a DNA sequence is substantially homologous" refers with respect to the crtE encoding DNA sequence to a DNA sequence which encodes an amino acid sequence which shows more than 45 %, preferably more than 60 % and more preferably more than 75 % and most preferably more than 90 % identical amino acids when compared to the amino acid sequence of crtE of Flavobacterium sp. 1534 and is the amino acid sequence of a polypeptide which shows the same type of enzymatic activity as the enzyme encoded by crtE of Flavobacterium sp. 1534. In analogy with respect to crtB this means more than 60 %, preferably more than 70 %, more preferably more than 80 % and most preferably more than 90 %; with respect to crtI this means more than 70 %, preferably more than 80 % and most preferably more than 90 %; with respect to crtY this means 55 %, preferably 70 %, more preferably 80 % and most preferably 90 %.

"DNA sequences which are substantially homologous" refer with respect to the $crtW_{E396}$ encoding DNA sequence to a DNA sequence which encodes an amino acid sequence which shows more than 60%, preferably more than 75% and most preferably more than 90% identical amino acids when compared to the amino acid sequence of crtW of the microorganism E 396 (FERM BP-4283) and is the amino acid sequence of a polypeptide which shows the same type of enzymatic activity as the enzyme encoded by crtW of the microorganism E 396. In analogy with respect to $crtZ_{E396}$ this means more than 75%, preferable more than 80% and most preferably more than 90%; with respect to $crtE_{E396}$, $crtB_{E396}$, $crtI_{E396}$, $crtY_{E396}$ and $crtZ_{E396}$ this means more than 80%, preferably more than 90% and most preferably 95%.

The expression "said polynucleotide being substantially free of other polynucleotides of Flavobacterium sp. R1534" and "said polynucleotide being substantially free of other polynucleotides of microorganism E-396" is meant to preclude the present invention from encompassing the polynucleotides as they exist in Flavobacterium sp. R1534 or in microorganism E-396, themselves. The polynucleotides herein described which are combinations of two or more DNA sequences of Flavobacterium sp. R1534 and/or microorganism E-396 are also substantially free of other polynucleotides of Flavobacterium sp. R1534 and microorganism E-396 in any circumstance where a polynucleotide containing only a single such DNA sequence would be substantially free of other polynucleotides of Flavobacterium sp. R1534 or microorganism E-396.

DNA sequences in form of genomic DNA, cDNA or synthetic DNA can be prepared as known in the art [see e.g. Sambrook et al., Molecular Cloning, Cold Spring Habor Laboratory Press 1989] or, e.g. as specifically described in Examples 1, 2 or 7. In the context of the present invention it should be noted that all DNA sequences used for the process for production of carotenoids of the present invention encoding crt-gene products can also be prepared as synthetic DNA sequences according to known methods or in analogy to the method specifically described for crtW in Example 7.

The cloning of the DNA-sequences of the present invention from such genomic DNA can than be effected, e.g. by using the well known polymerase chain reaction (PCR) method. The principles of this method are outlined e.g. in PCR Protocols: A guide to Methods and Applications, Academic Press, Inc. (1990). PCR is an in vitro method for producing large amounts of a specific DNA of defined

length and sequence from a mixture of different DNA-sequences. Thereby, PCR is based on the enzymatic amplification of the specific DNA fragment of interest which is flanked by two oligonucleotide primers which are specific for this sequence and which hybridize to the opposite strand of the target sequence. The primers are oriented with their 3' ends pointing toward each other. Repeated cycles of heat denaturation of the template, annealing of the primers to their complementary sequences and extension of the annealed primers with a DNA polymerase result in the amplification of the segment between the PCR primers. Since the extension product of each primer can serve as a template for the other, each cycle essentially doubles the amount of the DNA fragment produced in the previous cycle.

By utilizing the thermostable Taq DNA polymerase, isolated from the thermophilic bacteria Thermus aquaticus, it has been possible to avoid denaturation of the polymerase which necessitated the addition of enzyme after each heat denaturation step. This development has led to the automation of PCR by a variety of simple temperature-cycling devices. In addition, the specificity of the amplification reaction is increased by allowing the use of higher temperatures for primer annealing and extension. The increased specificity improves the overall yield of amplified products by minimizing the competition by non-target fragments for enzyme and primers. In this way the specific sequence of interest is highly amplified and can be easily separated from the non-specific sequences by methods known in the art, e.g. by separation on an agarose gel and cloned by methods known in the art using vectors as described e.g. by Holten and Graham in Nucleic Acid Res. 19, 1156 (1991), Kovalic et. al. in Nucleic Acid Res. 19, 4560 (1991), Marchuk et al. in Nucleic Acid Res. 19, 1154 (1991) or Mead et al. in Bio/Technology 9, 657-663 (1991).

The oligonucleotide primers used in the PCR procedure can be prepared as known in the art and described e.g. in Sambrook et al., s.a.

Amplified DNA-sequences can than be used to screen DNA libraries by methods known in the art (Sambrook et al., s.a.) or as specifically described in Examples 1 and 2.

Once complete DNA-sequences of the present invention have been obtained they can be used as a guideline to define new PCR primers for the cloning of substantially homologous DNA sequences from other sources. In addition they

and such homologous DNA sequences can be integrated into vectors by methods known in the art and described, e.g., in Sambrook et al. (s.a.) to express or overexpress the encoded polypeptide(s) in appropriate host systems. The expression vector into which the polynucleotides of the invention are integrated is not critical. Conventional expression vectors may be selected based upon the size of the polynucleotide of the invention to be inserted into the vector and the host cell to be transformed by the vector. Such conventional expression vectors contain a regulatory sequence for the synthesis of mRNA derived from the polynucleotide of the invention being expressed and possible marker genes. Conventional regulatory sequences generally contain, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence.

However, a man skilled in the art knows that also the DNA-sequences themselves can be used to transform the suitable host systems of the invention to get overexpression of the encoded polypeptide. Appropriate host systems are for example Bacteria e.g. E. coli, Bacilli as, e.g. Bacillus subtilis or Flavobacter strains. E. coli, which could be used are E. coli K12 strains e.g. M15 [described as DZ 291 by Villarejo et al. in J. Bacteriol. 120, 466-474 (1974)], HB 101 [ATCC No. 33694] or E. coli SG13009 [Gottesman et al., J. Bacteriol. 148, 265-273 (1981)].

Suitable Flavobacter strains can be obtained from any of the culture collections known to the man skilled in the art and listed, e.g. in the journal "Industrial Property" (January 1994, pgs 29-40), like the American Type Culture Collection (ATCC) or the Centralbureau voor Schimmelkultures (CBS) and are, e.g. Flavobacterium sp. R 1534 (ATCC No. 21588, classified as unknown bacterium; or as CBS 519.67) or all Flavobacter strains listed as CBS 517.67 to CBS 521.67 and CBS 523.67 to CBS 525.67, especially R 1533 (which is CBS 523.67 or ATCC 21081, classified as unknown bacterium; see also USP 3,841,967). Further Flavobacter strains are also described in WO 91/03571. Suitable eukaryotic host systems are for example fungi, like Aspergilli, e.g. Aspergillus niger [ATCC 9142] or yeasts, like Saccharomyces, e.g. Saccharomyces cerevisiae or Pichia, like pastoris, all available from ATCC.

Suitable vectors which can be used for expression in E. coli are mentioned, e.g., by Sambrook et al. [s.a.] or by Fiers et al. in Procd. 8th Int. Biotechnology Symposium" [Soc. Franc. de Microbiol., Paris (Durand et al., eds.), pp. 680-697

(1988)] or by Bujard et al. in Methods in Enzymology, eds. Wu and Grossmann, Academic Press, Inc. Vol. 155, 416-433 (1987) and Stüber et al. in Immunological Methods, eds. Lefkovits and Pernis, Academic Press, Inc., Vol. IV, 121-152 (1990). Vectors which could be used for expression in Bacilli are known in the art and described, e.g., in EP 405 370, EP 635 572 Procd. Nat. Acad. Sci. USA 81, 439 (1984) by Yansura and Henner, Meth. Enzym. 185, 199-228 (1990) or EP 207 459. Vectors which can be used for expression in fungi are known in the art and described e.g. in EP 420 358 and for yeast in EP 183 070, EP 183 071, EP 248 227, EP 263 311. Vectors which can be used for expression in Flavobacter are known in the art and described in the Examples or, e.g. in Plasmid Technology, edt. by J. Grinsted and P.M. Bennett, Academic Press (1990).

Once such DNA-sequences have been expressed in an appropriate host cell in a suitable medium, the carotenoids can be isolated either from the medium in the case they are secreted into the medium or from the host organism and, if necessary separated from other carotenoids if present in case one specific carotenoid is desired by methods known in the art (see e.g. Carotenoids Vol IA: Isolation and Analysis, G. Britton, S. Liaaen-Jensen, H. Pfander; 1995, Birkhäuser Verlag, Basel) .

The carotenoids of the present invention can be used in a process for the preparation of food or feeds. A man skilled in the art is familiar with such processes. Such compound foods or feeds can further comprise additives or components generally used for such purpose and known in the state of the art.

After the invention has been described in general hereinbefore, the following examples are intended to illustrate details of the invention, without thereby limiting it in any matter.

Example 1

Materials and general methods used

Bacterial strains and plasmids: *Flavobacterium sp.* R1534 WT (ATCC 21588) was the DNA source for the genes cloned. Partial genomic libraries of *Flavobacterium sp.* R1534 WT DNA were constructed into the pBluescriptII+(KS) or (SK) vector (Stratagene, La Jolla, USA) and transformed into *E. coli* XL-1 blue (Stratagene) or JM109.

Media and growth conditions: Transformed *E. coli* were grown in Luria broth (LB) at 37° C with 100mg Ampicillin (Amp)/ml for selection. *Flavobacterium sp.* R1534 WT was grown at 27° C in medium containing 1% glucose, 1% tryptone (Difco Laboratories), 1% yeast extract (Difco), 0.5% $MgSO_4 7H_2O$ and 3% NaCl.

Colony screening: Screening of the *E. coli* transformants was done by PCR basically according to the method described by Zon et al. [Zon et al., BioTechniques 7, 696-698 (1989)] using the following primers:

Primer #7 (SEQ ID NO: 38): 5'-CCTGGATGACGTGCTGGAATATTCC-3'

Primer #8 (SEQ ID NO: 39): 5'-CAAGGCCCAGATCGCAGGCG-3'

Genomic DNA: A 50 ml overnight culture of *Flavobacterium sp. R1534* was centrifuged at 10,000 g for 10 minutes. The pellet was washed briefly with 10 ml of lysis buffer (50 mM EDTA, 0.1M NaCl pH7.5), resuspended in 4 ml of the same buffer sumplemented with 10 mg of lysozyme and incubated at 37°C for 15 minutes. After addition of 0.3 ml of N-Lauroyl sarcosine (20%) the incubation at 37°C was continued for another 15 minutes before the extraction of the DNA with phenol, phenol/chloroform and chloroform. The DNA was ethanol precipitated at room temperature for 20 minutes in the presence of 0.3 M sodium acetate (pH 5.2), followed by centrifugation at 10,000 g for 15 minutes. The pellet was rinsed with 70% ethanol, dried and resuspended in 1 ml of TE (10 mM Tris, 1mM EDTA, pH 8.0).

All genomic DNA used in the southern blot analysis and cloning experiments was dialysed against H_2O for 48 hours, using collodium bags (Sartorius, Germany), ethanol precipitated in the presence of 0.3 M sodium acetate and resuspended in H_2O.

Probe labelling: DNA probes were labeled with (a - ^{32}P) dGTP (Amersham) by random-priming according to [Sambrook et al., s.a.].

Probes used to screen the mini-libraries: Probe 46F is a 119 bp fragment obtained by PCR using primer #7 (SEQ ID NO: 38) and #8 (SEQ ID NO: 39) and *Flavobacterium sp.* R1534 genomic DNA as template. This probe was proposed to be a fragment of the *Flavobacterium sp.* R1534 phytoene synthase (crtB) gene, since it shows significant homology to the phytoene synthase genes from other species (e.g.

E. uredovora, E. herbicola). **Probe A** is a BstXI - PstI fragment of 184 bp originating from the right arm of the insert of clone 85. **Probe B** is a 397 bp XhoI - NotI fragment obtained from the left end of the insert of clone 85. **Probe C** is a 536 bp BglII - PstI fragment from the right end of the insert of clone 85. **Probe D** is a 376 bp KpnI - BstYI fragment isolated from the insert of clone 59. The localization of the individual probes is shown in figure 6.

Oligonucleotide synthesis: The oligonucleotides used for PCR reactions or for sequencing were synthesized with an Applied Biosystems 392 DNA synthesizer.

Southern blot analysis: For hybridization experiments *Flavobacterium sp. R1534* genomic DNA (3 mg) was digested with the appropiate restriction enzymes and electrophoresed on a 0.75% agarose gel. The transfer to Zeta-Probe blotting membranes (BIO-RAD), was done as described [Sourthern, E.M., J. Mol. Biol. 98, 503 (1975)]. Prehybridization and hybridization was in 7%SDS, 1% BSA (fraction V; Boehringer), 0.5M Na_2HPO_4, pH 7.2 at 65°C. After hybridization the membranes were washed twice for 5 minutes in 2x SSC, 1% SDS at room temperature and twice for 15 minutes in 0.1% SSC, 0.1% SDS at 65° C.

DNA sequence analysis: The sequence was determined by the dideoxy chain termination technique [Sanger et al., Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977)] using the Sequenase Kit (United States Biochemical). Both strands were completely sequenced and the sequence analyzed using the GCG sequence analysis software package (Version 8.0) by Genetics Computer, Inc. [Devereux et al., Nucleic Acids. Res. 12, 387-395 (1984)].

Analysis of carotenoids: *E. coli* XL-1 or JM109 cells (200 - 400 ml) carrying different plasmid constructs were grown for the times indicated in the text, usually 24 to 60 hours, in LB suplemented with 100mg Ampicillin/ml, in shake flasks at 37° C and 220 rpm.

The carotenoids present in the microorganisms were extracted with an adequate volume of acetone using a rotation homogenizer (Polytron, Kinematica AG, CH-Luzern). The homogenate was the filtered through the sintered glass of a suction filter into a round bottom flask. The filtrate was evaporated by means of a rotation evaporator at 50° C using a water-jet vacuum. For the zeaxanthin detection the residue was dissolved in n-hexane/acetone (86:14) before analysis with a normalphase HPLC as described in [Weber, S. in Analytical Methods for

Vitamins and Carotenoids in Feed, Keller, H.E. , Editor, 83-85 (1988)]. For the detection of β-carotene and lycopene the evaporated extract was dissolved in n-hexane/acetone (99:1) and analysed by HPLC as described in [Hengartner et al., Helv. Chim. Acta 75, 1848-1865 (1992)].

Example 2

Cloning of the *Flavobacterium sp.* R1534 carotenoid biosynthetic genes.

To identify and isolate DNA fragments carrying the genes of the carotenoid biosynthesis pathway, we used the DNA fragment 46F (see methods) to probe a Southern Blot carrying chromosomal DNA of *Flavobacterium sp.* R1534 digested with different restriction enzymes Fig. 2. The 2.4 kb XhoI/PstI fragment hybridizing to the probe seemed the most appropiate one to start with. Genomic *Flavobacterium sp.* R1534 DNA was digested with XhoI/PstI and run on a 1% agarose gel. According to a comigrating DNA marker, the region of about 2.4 kb was cut out of the gel and the DNA isolated. A XhoI/PstI mini library of *Flavobacterium sp.* R1534 genomic DNA was constructed into XhoI - PstI sites of pBluescriptIISK(+). One hundred *E. coli* XL1 transformants were subsequently screened by PCR with primer #7 (SEQ ID NO: 38) and primer # 8 (SEQ ID NO: 39), the same primers previously used to obtain the 119 bp fragment (46F). One positive transformant, named clone 85, was found. Sequencing of the insert revealed sequences not only homologous to the phytoene synthase (crtB) but also to the phytoene desaturase (crtI) of both *Erwinia* species herbicola and uredovora. Left and right hand genomic sequences of clone 85 were obtained by the same approach using probe A and probe B. *Flavobacterium sp.* R1534 genomic DNA was double digested with ClaI and Hind III and subjected to Southern analysis with probe A and probe B. With probe A a ClaI/HindIII fragment of aprox. 1.8 kb was identified (Fig. 3A), isolated and subcloned into the ClaI/HindIII sites of pBluescriptIIKS (+) . Screening of the *E. coli* XL1 transformants with probe A gave 6 positive clones. The insert of one of these positives, clone 43-3, was sequenced and showed homology to the N-terminus of crtI genes and to the C-terminus of crtY genes of both *Erwinia* species mentioned above. With probe B an approx. 9.2 kb ClaI/HindIII fragment was detected (Fig. 3B), isolated and subcloned into pBluescriptIIKS (+) .

A screening of the transformants gave one positive, clone 51. Sequencing of the 5' and 3' of the insert, revealed that only the region close to the HindIII site showed relevant homology to genes of the carotenoid biosynthesis of the *Erwinia*

species mentioned above (e.g. crtB gene and crtE gene). The sequence around the ClaI site showed no homology to known genes of the carotenoid biosynthesis pathway. Based on this information and to facilitate further sequencing and construction work, the 4.2 kb BamHI/HindIII fragment of clone 51 was subcloned into the respective sites of pBluescriptIIKS(+) resulting in clone 2. Sequencing of the insert of this clone confirmed the presence of genes homologous to *Erwinia sp* crtB and crtE genes. These genes were located within 1.8 kb from the HindIII site. The remaining 2.4 kb of this insert had no homology to known carotenoid biosynthesis genes.

Additional genomic sequences downstream of the ClaI site were detected using probe C to hybridize to *Flavobacterium sp.* R1534 genomic DNA digested with different restriction enzymes (see figure 4).

A SalI/HindIII fragment of 2.8 kb identified by Southern analysis was isolated and subcloned into the HindIII/XhoI sites of pBluescriptIIKS (+). Screening of the *E. coli* XL1 transformants with probe A gave one positive clone named clone 59. The insert of this clone confirmed the sequence of clone 43-3 and contained in addition sequences homologous to the N-terminus of the crtY gene from other known lycopene cyclases. To obtain the putative missing crtZ gene a Sau3AI partial digestion library of *Flavobacterium sp. R1534* was constructed into the BamHI site of pBluescriptIIKS(+). Screening of this library with probe D gave several positive clones. One transformant designated, clone 6a, had an insert of 4.9 kb. Sequencing of the insert revealed besides the already known sequences coding for crtB, crtI and crtY also the missing crtZ gene. Clone 7g was isolated from a mini library carrying BclI/SphI fragments of R1534 (Fig. 5) and screened with probe D. The insert size of clone 7g is approx. 3 kb.

The six independent inserts of the clones described above covering approx. 14 kb of the *Flavobacterium sp.* R1534 genome are compiled in Figure 6.

The determined sequence spanning from the BamHI site (position 1) to base pair 8625 is shown figure 7.

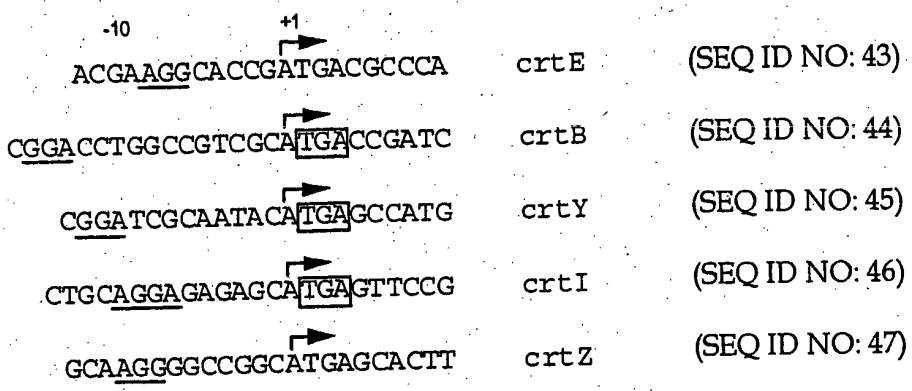
Putative protein coding regions of the cloned R1534 sequence.

Computer analysis using the CodonPreference program of the GCG package, which recognizes protein coding regions by virtue of the similarity of their codon usage to a given codon frequency table, revealed eight open reading frames (ORFs)

encoding putative proteins: a partial ORF from 1 to 1165 (ORF-5) (SEQ ID NO: 41) coding for a polypeptide larger than 41382 Da; an ORF coding for a polypeptide with a molecular weight of 40081 Da from 1180 to 2352 (ORF-1) (SEQ ID NO: 40); an ORF coding for a polypeptide with a molecular weight of 31331 Da from 2521 to 3405 (crtE); an ORF coding for a polypeptide with a molecular weight of 32615 Da from 4316 to 3408 (crtB); an ORF coding for a polypeptide with a molecular weight of 54411 Da from 5797 to 4316 (crtI); an ORF coding for a polypeptide with a molecular weight of 42368 Da from 6942 to 5797 (crtY); an ORF coding for a polypeptide with a molecular weight of 19282 Da from 7448 to 6942 (crtZ); and an ORF coding for a polypeptide with a molecular weight of 19368 Da from 8315 to 7770 (ORF-16) (SEQ ID NO: 42); ORF-1 and crtE have the opposite transcriptional orientation from the others (Fig. 6). The translation start sites of the ORFs crtI, crtY and crtZ could clearly be determined based on the appropiately located sequences homologous to the Shine/Delgano (S/D) [Shine and Dalgarno, Proc. Natl. Acad. Sci. USA 71, 1342-1346 (1974)] consensus sequence AGG--6-9N--ATG (Fig. 10) and the homology to the N-terminal sequences of the respective enzymes of *E. herbicola* and *E. uredovora*. The translation of the ORF crtB could potentially start from three closely spaced codons ATG (4316), ATG (4241) and ATG (4211). The first one, although not having the best S/D sequence of the three, gives a translation product with the highest homology to the N-terminus of the *E. herbicola* and *E. uredovora* crtB protein, and is therefore the most likely translation start site. The translation of ORF crtE could potentially start from five different start codons found within 150 bp : ATG (2389), ATG (2446), ATG (2473), ATG (2497) and ATG (2521). We believe that based on the following observations, the ATG (2521) is the most likely transcription start site of crtE: this ATG start codon is preceeded by the best consensus S/D sequence of all five putative start sites mentioned; and the putative N-terminal amino acid sequence of the protein encoded has the highest homology to the N-terminus of the crtE enzymes of *E. herbicola* and *E. uredovora*;

Characteristics of the crt translational initiation sites and gene products.

The translational start sites of the five carotenoid biosynthesis genes are shown below and the possible ribosome binding sites are underlined. The genes crtZ, crtY, crtI and crtB are grouped so tightly that the TGA stop codon of the anterior gene overlaps the ATG of the following gene. Only three of the five genes (crtI, crtY and crtZ) fit with the consensus for optimal S/D sequences. The boxed TGA sequence shows the stop condon of the anterior gene.



Amino acid sequences of individual crt genes of *Flavobacterium sp.* R1534.

All five ORFs of *Flavobacterium sp.* R1534 having homology to known carotenoid biosynthesis genes of other species are clustered in approx. 5.2 kb of the sequence (Fig. 7) (SEQ ID NO: 1).

GGDP synthase (crtE)

The amino acid (aa) sequence of the geranylgeranyl pyrophosphate synthase (crtE gene product) consists of 295 aa and is shown in figure 8 (SEQ ID NO: 2). This enzyme condenses farnesyl pyrophosphate and isopentenyl pyrophosphate in a 1' - 4.

Phytoene synthase (crtB)

This enzyme catalyzes two enzymatic steps. First it condenses in a head to head reaction two geranylgeranyl pyrophosphates (C20) to the C40 carotenoid prephytoene. Second it rearanges the cyclopropylring of prephytoene to phytoene. The 303 aa encoded by the crtB gene of *Flavobacterium sp.* R1534 is shown in figure 9 (SEQ ID NO: 3).

Phytoene desaturase (crtI)

The phytoene desaturase of *Flavobacterium sp.* R1534 consisting of 494 aa, shown in figure 10 (SEQ ID NO: 4), performs like the crtI enzyme of *E. herbicola* and *E. uredovora*, four desaturation steps, converting the non-coloured carotenoid phytoene to the red coloured lycopene.

Lycopene cyclase (crtY)

The crtY gene product of *Flavobacterium sp.* R1534 is sufficient to introduce the b-ionone rings at both sides of lycopene to obtain β-carotene. The lycopene cyclase of *Flavobacterium sp.* R1534 consists of 382 aa (Fig. 11) (SEQ ID NO: 5).

β-carotene hydroxylase (crtZ)

The gene product of crtZ consisting of 169 aa (Fig. 12) (SEQ ID NO: 6) and hydroxylates β-carotene to the xanthophyll zeaxanthin.

Putative enzymatic functions of the ORF's (orf-1 (SEQ ID NO: 40), **orf-5** (SEQ ID NO: 41) **and orf-16** (SEQ ID NO: 42))

The orf-1 (SEQ ID NO: 40) has at the aa level over 40% identity to acetoacetyl-CoA thiolases of different organisms (e.g. Candida tropicalis, human, rat). This gene is therefore most likely a putative acetoacetyl-CoA thiolase (acetyl-CoA acetyltransferase), which condenses two molecules of acetyl-CoA to Acetoacetyl-CoA. Condensation of acetoacetyl-CoA with a third acetyl-CoA by the HMG-CoA synthase forms β-hydroxy-β-methylglutaryl-CoA (HMG-CoA). This compound is part of the mevalonate pathway which produces besides sterols also numerous kinds of isoprenoids with diverse cellular functions. In bacteria and plants, the isoprenoid pathway is also able to synthesize some unique products like carotenoids, growth regulators (e.g. in plants gibberellins and abcissic acid) and sencodary metabolites like phytoalexins [Riou et al., Gene 148, 293-297 (1994)].

The orf-5 (SEQ ID NO: 41) has a low homology of approx. 30%, to the amino acid sequence of polyketide synthases from different streptomyces (e.g. S. violaceoruber, S. cinnamonensis). These antibiotic synthesizing enzymes (polyketide synthases), have been classified into two groups. Type-I polyketide synthases are large multifunctional proteins, whereas type-II polyketide synthases are multiprotein complexes composed of several individual proteins involved in the subreactions of the polyketide synthesis [Bibb, et al. Gene 142, 31-39 (1994)].

The putative protein encoded by the orf-16 (SEQ ID NO: 42) has at the aa level an identity of 42% when compared to the soluble hydrogenase subunit of Anabaena cylindrica.

Functional assignment of the ORF 's (crtE, crtB, crtI, crtY and crtZ) to

enzymatic activities of the carotenoid biosynthesis pathway.

The biochemical assignment of the gene products of the different ORF's were revealed by analyzing carotenoid accumulation in *E. coli* host strains that were transformed with deleted variants of the *Flavobacterium sp.* gene cluster and thus expressed not all of the crt genes (Fig. 13).

Three different plasmid were constructed: pLyco, p59-2 and pZea4. Plasmid p59-2 was obtained by subcloning the HindIII/BamHI fragment of clone 2 into the HindIII/BamHI sites of clone 59. p59-2 carries the ORF's of the crtE, crtB, crtI and crtY gene and should lead to the production of β-carotene. pLyco was obtained by deleting the KpnI/KpnI fragment, coding for approx. one half (N-terminus) of the crtY gene, from the p59-2 plasmid. *E. coli* cells transformed with pLyco, and therefore having a truncated non-functional crtY gene, should produce lycopene, the precursor of β-carotene. pZea4 was constructed by ligation of the AscI-SpeI fragment of p59-2, containing the crtE, crtB, crtI and most of the crtY gene with the AscI/XbaI fragment of clone 6a, containing the sequences to complete the crtY gene and the crtZ gene. pZea4 [for complete sequence see Fig. 24 (SEQ ID NO: 27); nucleotides 1 to 683 result from pBluescriptIIKS(+), nucleotides 684 to 8961 from Flavobacterium R1534 WT genome, nucleotides 8962 to 11233 from pBluescriptIIKS(+)] has therefore all five ORF's of the zeaxanthin biosynthesis pathway. Plasmid pZea4 has been deposited on May 25, 1995 at the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Germany) under accession No. DSM 10012. *E. coli* cells transformed with this latter plasmid should therefore produce zeaxanthin. For the detection of the carotenoid produced, transformants were grown for 48 hours in shake flasks and then subjected to carotenoid analysis as described in the methods section. Figure 13 summarizes the different inserts of the plasmids described above, and the main carotenoid detected in the cells.

As expected the pLyco carrying *E. coli* cells produced lycopene, those carrying p59-2 produced β-carotene (all-E,9-Z,13-Z) and the cells having the pZea4 construct produced zeaxanthin. This confirms that all the necessary genes of *Flavobacterium sp.* R1534 for the synthesis of zeaxanthin or their precursors (phytoene, lycopene and β-carotene) were cloned.

Example 3

Materials and methods used for expression of carotenoid synthesizing enzymes

Bacterial strains and plasmids: The vectors pBluescriptIIKS (+) or (-) (Stratagene, La Jolla, USA) and pUC18 [Vieira and Messing, Gene 19, 259-268 (1982); Norrander et al., Gene 26, 101-106 (1983)] were used for cloning in different *E. coli* strains, like XL-1 blue (Stratagene), TG1 or JM109. In all *B. subtilis* transformations, strain 1012 was used. Plasmids pHP13 [Haima et al., Mol. Gen. Genet. 209, 335-342 (1987)] and p602/22 [LeGrice, S.F.J. in Gene Expression Technology, Goeddel, D.V., Editor, 201-214 (1990)] are Gram (+)/(-) shuttle vectors able to replicate in *B. subtilis* and *E. coli* cells. Plasmid p205 contains the vegI promoter cloned into the SmaI site of pUC18. Plasmid pXI12 is an integration vector for the constitutive expression of genes in *B. subtilis* [Haiker et al., in 7th Int. Symposium on the Genetics of Industrial Microorganisms, June 26-July 1, 1994. Montreal, Quebec, Canada (1994)]. Plasmid pBEST501 [Itaya et al., Nucleic Acids Res. 17 (11), 4410 (1989)] contains the neomycin resistance gene cassette originating from the plasmid pUB110 (GenBank entry: M19465) of *S. aureus* [McKenzie et al., Plasmid 15, 93-103 (1986); McKenzie et al., Plasmid 17, 83-84 (1987)]. This neomycin gene has been shown to work as a selection marker when present in a single copy in the genome of *B. subtilis*. Plasmid pC194 (ATCC 37034)(GenBank entry: L08860) originates from *S. aureus* [Horinouchi and Weisblaum, J. Bacteriol. 150, 815-825 (1982)] and contains the chloramphenicol acetyltransferase gene.

Media and growth conditions: *E. coli* were grown in Luria broth (LB) at 37° C with 100mg Ampicillin (Amp)/ml for selection. *B. subtilis* cells were grown in VY-medium supplemented with either erythromycin (1 mg/ml), neomycin (5-180 mg/ml) or chloramphenicol (10-80 mg/ml).

Transformation: *E. coli* transformations were done by electroporation using the Gene-pulser device of BIO-RAD (Hercules, CA, USA) with the following parameters (200 W, 250 mFD, 2.5V). *B.subtilis* transformations were done basically according to the standard procedure method 2.8 described by [Cutting and Vander Horn in Molecular Biological Methods for Bacillus, Harwood, C.R. and Cutting, S.M., Editor, John Wiley & Sons: Chichester, England. 61-74 (1990)].

Colony screening: Bacterial colony screening was done as described by [Zon et al., s.a.].

Oligonucleotide synthesis: The oligonucleotides used for PCR reactions or for sequencing were synthesized with an Applied Biosystems 392 DNA synthesizer.

PCR reactions: The PCR reactions were performed using either the **UlTma DNA polymerase** (Perkin Elmer Cetus) or the **Pfu Vent polymerase** (New England Biolabs) according to the manufacturers instructions. A typical 50 ml PCR reaction contained: 100ng of template DNA, 10 pM of each of the primers, all four dNTP's (final conc. 300 mM), $MgCl_2$ (when UlTma polymerase was used; final conc. 2 mM), 1x UlTma reaction buffer or 1x Pfu buffer (supplied by the manufacturer). All components of the reaction with the exception of the DNA polymerase were incubated at 95°C for 2 min. followed by the cycles indicated in the respective section (see below). In all reactions a hot start was made, by adding the polymerase in the first round of the cycle during the 72°C elongation step. At the end of the PCR reaction an aliquot was analysed on 1% agarose gel, before extracting once with phenol/chloroform. The amplified fragment in the aqueous phase was precipitated with 1/10 of a 3M NaAcetate solution and two volumes of Ethanol. After centrifugation for 5 min. at 12000 rpm, the pellet was resuspended in an adequate volume of H_2O, typically 40 ml, before digestion with the indicated restriction enzymes was performed. After the digestion the mixture was separated on a 1% low melting point agarose. The PCR product of the expected size were excised from the agarose and purified using the glass beads method (GENECLEAN KIT, Bio 101, Vista CA, USA) when the fragments were above 400 bp or directly spun out of the gel when the fragments were shorter than 400 bp, as described by [Heery et al., TIBS 6 (6), 173 (1990)].

Oligos used for gene amplification and site directed mutagenesis:

All PCR reactions performed to allow the construction of the different plasmids are described below. All the primers used are summarized in figure 14.

Primers #100 (SEQ ID NO: 7) and **#101** (SEQ ID NO: 8) were used in a PCR reaction to amplify the complete crtE gene having a SpeI restriction site and an artificial ribosomal binding site (RBS) upstream of the transcription start site of this gene. At the 3' end of the amplified fragment, two unique restriction sites were introduced, an AvrII and a SmaI site, to facilitate the further cloning steps. The

PCR reaction was done with UlTma polymerase using the following conditions for the amplification: 5 cycles with the profile: 95°C, 1 min./ 60°C, 45 sec./ 72°C, 1 min. and 20 cycles with the profile: 95°C, 1 min./ 72°C, 1 min.. Plasmid pBIIKS(+)-clone2 served as template DNA. The final PCR product was digested with SpeI and SmaI and isolated using the GENECLEAN KIT. The size of the fragment was approx. 910 bp.

Primers **#104** (SEQ ID NO: 9) and **#105** (SEQ ID NO: 10) were used in a PCR reaction to amplify the crtZ gene from the translation start till the SalI restriction site, located in the coding sequence of this gene. At the 5' end of the crtZ gene an EcoRI, a synthetic RBS and a NdeI site was introduced. The PCR conditions were as described above. Plasmid pBIIKS(+)-clone 6a served as template DNA and the final PCR product was digested with EcoRI and SalI. Isolation of the fragment of approx. 480 bp was done with the GENECLEAN KIT.

Primers **MUT1** (SEQ ID NO: 11) and **MUT5** (SEQ ID NO: 14) were used to amplify the complete crtY gene. At the 5' end, the last 23 nucleotides of the crtZ gene including the SalI site are present, followed by an artificial RBS preceding the translation start site of the crtY gene. The artificial RBS created includes a PmlI restriction site. The 3' end of the amplified fragment contains 22 nucleotides of the crtI gene, preceded by a newly created artifial RBS which contains a MunI restriction site. The conditions used for the PCR reaction were as described above using the following cycling profile: 5 rounds of 95°C, 45 sec./ 60°C, 45 sec./ 72°C, 75 sec. followed by 22 cycles with the profile: 95°C, 45 sec./ 66°C, 45 sec./ 72°C, 75 sec.. Plasmid pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. The PCR product of 1225 bp was made blunt and cloned into the SmaI site of pUC18, using the Sure-Clone Kit (Pharmacia) according to the manufacturer.

Primers **MUT2** (SEQ ID NO: 15) and **MUT6** (SEQ ID NO: 15) were used to amplify the complete crtI gene. At the 5' the last 23 nucleotides of the crtY gene are present, followed by an artificial RBS which precedes the translation start site of the crtI gene. The new RBS created, includes a MunI restriction site. The 3' end of the amplified fragment contains the artificial RBS upstream of the crtB gene including a BamHI restriction site. The conditions used for the PCR reaction were basically as described above including the following cycling profile: 5 rounds of 95°C, 30 sec./ 60°C, 30 sec./ 72°C, 75 sec., followed by 25 cycles with the profile: 95°C, 30 sec./ 66°C, 30 sec./ 72°C, 75 sec.. Plasmid pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. For the further cloning steps the PCR product of 1541 bp was

digested with MunI and BamHI.

Primers **MUT3** (SEQ ID NO: 13) and **CAR17** (SEQ ID NO: 16) were used to amplify the N-terminus of the crtB gene. At the 5' the last 28 nucleotides of the crtI gene are present followed by an artificial RBS, preceding the translation start site of the crtB gene. This new created RBS, includes a BamHI restriction site. The amplified fragment, named PCR-F contains also the HindIII restriction site located at the N-terminus of the crtB gene. The conditions used for the PCR reaction were as described elsewhere in the text, including the following cycling profile: 5 rounds of 95°C, 30 sec./ 58°C, 30 sec./ 72°C, 20 sec. followed by 25 cycles with the profile: 95°C, 30 sec./ 60°C, 30 sec./ 72°C, 20 sec.. Plasmid pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. The PCR product of approx. 160 bp was digested with BamHI and HindIII.

Oligos used to amplify the chloramphenicol resistance gene (cat):

Primers **CAT3** (SEQ ID NO: 17) and **CAT4** (SEQ ID NO: 18) were used to amplify the chloramphenicol resistance gene of pC194 (ATCC 37034) [Horinouchi and Weisblum, s.a.] a R-plasmid found in *S. aureus*. The conditions used for the PCR reaction were as described previously including the following cycling profile: 5 rounds of 95°C, 60 sec./ 50°C, 60 sec./ 72°C, 2 min. followed by 20 cycles with the profile: 95°C, 60 sec./ 60°C, 60 sec./ 72°C, 2 min.. Plasmid pC198 served as template for the Pfu Vent polymerase. The PCR product of approx. 1050 bp was digested with EcoRI and
AatII.

Oligos used to generate linkers: Linkers were obtained by adding 90 ng of each of the two corresponding primers into an Eppendorf tube. The mixture was dried in a speed vac and the pellet resuspended in 1x Ligation buffer (Boehringer, Mannheim, Germany). The solution was incubated at 50°C for 3 min. before cooling down to RT, to allow the primers to hybridize properly. The linker were now ready to be ligated into the appropriate sites. All the oligos used to generate linkers are shown in figure 15.

Primers **CS1** (SEQ ID NO: 19) and **CS2** (SEQ ID NO: 20) were used to form a linker containing the following restrictions sites HindIII, AflII, ScaI, XbaI, PmeI and EcoRI.

Primers **MUT7** (SEQ ID NO: 21) and **MUT8** (SEQ ID NO: 22) were used to

form a linker containing the restriction sites SalI, AvrII, PmlI, MluI, MunI, BamHI, SphI and HindIII.

Primers **MUT9** (SEQ ID NO: 23) and **MUT10** (SEQ ID NO: 24) were used to introduce an artificial RBS upstream of crtY.

Primers **MUT11** (SEQ ID NO: 25) and **MUT12** (SEQ ID NO: 26) were used to introduce an artificial RBS upstream of crtE.

Isolation of RNA: Total RNA was prepared from log phase growing *B. subtilis* according to the method described by [Maes and Messens, Nucleic Acids Res. 20 (16), 4374 (1992)].

Northern Blot analysis: For hybridization experiments up to 30 mg of *B. subtilis* RNA was electrophoreses on a 1% agarose gel made up in 1x MOPS and 0.66 M formaldehyde. Transfer to Zeta-Probe blotting membranes (BIO-RAD), UV cross-linking, pre-hybridization and hybridization was done as described elsewhere in [Farrell, J.R.E., RNA Methodologies. A laboratory Guide for isolation and characterization. San Diego, USA: Academic Press (1993)]. The washing conditions used were: 2 x 20 min. in 2xSSPE/0.1% SDS followed by 1 x 20 min. in 0.1% SSPE/0.1% SDS at 65°C. Northern blots were then analyzed either by a Phosphorimager (Molecular Dynamics) or by autoradiography on X-ray films from Kodak.

Isolation of genomic DNA: *B. subtilis* genomic DNA was isolated from 25 ml overnight cultures according to the standard procedure method 2.6 described by [13].

Southern blot analysis: For hybridization experiments *B. subtilis* genomic DNA (3 mg) was digested with the appropriate restriction enzymes and electrophoresed on a 0.75% agarose gel. The transfer to Zeta-Probe blotting membranes (BIO-RAD), was done as described [Southern, E.M., s.a.]. Prehybridization and hybridization was in 7%SDS, 1% BSA (fraction V; Boehringer), 0.5M Na_2HPO_4, pH 7.2 at 65°C. After hybridization the membranes were washed twice for 5 min. in 2x SSC, 1% SDS at room temperature and twice for 15 min. in 0.1% SSC, 0.1% SDS at 65° C. Southern blots were then analyzed either by a Phosphorimager (Molecular Dynamics) or by autoradiography on X-ray films from Kodak.

DNA sequence analysis: The sequence was determined by the dideoxy chain termination technique [Sanger et al., s.a.] using the Sequenase Kit Version 1.0 (United States Biochemical). Sequence analysis were done using the GCG sequence analysis software package (Version 8.0) by Genetics Computer, Inc. [Devereux et al., s.a.].

Gene amplification in *B. subtilis*: To amplify the copy number of the SFCO in *B. subtilis* transformants, a single colony was inoculated in 15 ml VY-medium supplemented with 1.5 % glucose and 0.02 mg chloramphenicol or neomycin/ml, dependend on the antibiotic resistance gene present in the amplifiable structure (see results and discussion). The next day 750 ml of this culture were used to inoculate 13 ml VY-medium containing 1.5% glucose supplemented with (60, 80, 120 and 150 mg/ml) for the cat resistant mutants, or 160 mg/ml and 180 mg/ml for the neomycin resistant mutants). The cultures were grown overnight and the next day 50 ml of different dilutions (1: 20, 1:400, 1: 8000, 1: 160'000) were plated on VY agar plates with the appropriate antibiotic concentration. Large single colonies were then further analyzed to determine the number of copies and the amount of carotenoids produced.

Analysis of carotenoids: *E. coli* or *B. subtilis* transformants (200 - 400 ml) were grown for the times indicated in the text, usually 24 to 72 hours, in LB-medium or VY-medium, respectively, supplemented with antibiotics, in shake flasks at 37° C and 220 rpm.

The carotenoids produced by the microorganisms were extracted with an adequate volume of acetone using a rotation homogenizer (Polytron, Kinematica AG, CH-Luzern). The homogenate was the filtered through the sintered glass of a suction filter into a round bottom flask. The filtrate was evaporated by means of a rotation evaporator at 50° C using a water-jet vacuum. For the zeaxanthin detection the residue was dissolved in n-hexane/acetone (86:14) before analysis with a normalphase HPLC as described in [Weber, S., s.a.]. For the detection of β-carotene and lycopene the evaporated extract was dissolved in n-hexane/acetone (99:1) and analysed by HPLC as described in Hengartner et al., s.a.].

Example 4

Carotenoid production in *E. coli*

The biochemical assignment of the gene products of the different open

reading frames (ORF's) of the carotenoid biosynthesis cluster of *Flavobacterium sp.* were revealed by analyzing the carotenoid accumulation in *E. coli* host strains, transformed with plasmids carrying deletions of the *Flavobacterium sp.* gene cluster, and thus lacking some of the crt gene products. Similar functional assays in *E. coli* have been described by other authors [Misawa et al., s.a.; Perry et al., J. Bacteriol., 168, 607-612 (1986); Hundle, et al., Molecular and General Genetics 254 (4), 406-416 (1994)]. Three different plasmid pLyco, pBIIKS(+)-clone59-2 and pZea4 were constructed from the three genomic isolates pBIIKS(+)-clone2, pBIIKS(+)-clone59 and pBIIKS(+)-clone6a (see figure 16).

Plasmid pBIIKS(+)-clone59-2 was obtained by subcloning the HindIII/BamHI fragment of pBIIKS(+)-clone 2 into the HindIII/BamHI sites of pBIIKS(+)-clone59. The resulting plasmid pBIIKS(+)-clone59-2 carries the complete ORF's of the crtE, crtB, crtI and crtY gene and should lead to the production of β-carotene. pLyco was obtained by deleting the KpnI/KpnI fragment, coding for approx. one half (N-terminus) of the crtY gene, from the plasmid pBIIKS(+)-clone59-2. *E. coli* cells transformed with pLyco, and therefore having a truncated non-functional crtY gene, should produce lycopene, the precursor of β-carotene. pZea4 was constructed by ligation of the AscI-SpeI fragment of pBIIKS(+)-clone59-2, containing the crtE, crtB, crtI and most of the crtY gene with the AscI/XbaI fragment of clone 6a, containing the crtZ gene and sequences to complete the truncated crtY gene mentioned above. pZea4 has therefore all five ORF's of the zeaxanthin biosynthesis pathway. *E. coli* cells transformed with this latter plasmid should therefore produce zeaxanthin. For the detection of the carotenoid produced, transformants were grown for 43 hours in shake flasks and then subjected to carotenoid analysis as described in the methods section. Figure 16 summarizes the construction of the plasmids described above.

As expected the pLyco carrying *E. coli* cells produced lycopene, those carrying pBIIKS(+)-clone59-2 produced β-carotene (all-E,9-Z,13-Z) and the cells having the pZea4 construct produced zeaxanthin. This confirms that we have cloned all the necessary genes of *Flavobacterium sp.* R1534 for the synthesis of zeaxanthin or their precursors (phytoene, lycopene and β-carotene). The production levels obtained are shown in table 1.

plasmid	host	zeaxanthin	β-χαροτενε	lycopene
pLyco	*E. coli JM109*	ND	ND	0.05%
pBIIKS(+)-clone59-2	"	ND	0.03%	ND
pZea4	"	0.033%	0.0009%	ND

Table 1: Carotenoid content of *E. coli* transformants, carrying the plasmids pLyco, pBIIKS(+)-clone59-2 and pZea4, after 43 hours of culture in shake flasks. The values indicated show the carotenoid content in % of the total dry cell mass (200 ml). ND = not detectable.

Examples 5

Carotenoid production in *B. subtilis*

In a first approach to produce carotenoids in *B. subtilis*, we cloned the carotenoid biosynthesis genes of *Flavobacterium* into the Gram (+)/(-) shuttle vectors p602/22, a derivative of p602/20 [LeGrice, S.F.J., s.a.]. The assembling of the final construct p602-CARVEG-E, begins with a triple ligation of fragments PvuII-AvrII of pZea4(del654-3028) and the AvrII-EcoRI fragment from plasmid pBIIKS(+)-clone6a, into the EcoRI and ScaI sites of the vector p602/22. The plasmid pZea4(del654-3028) had been obtained by digesting pZea4 with SacI and EspI. The protruding and recessed ends were made blunt with Klenow enzyme and religated. Construct pZea4(del654-3028) lacks most of the sequence upstream of crtE gene, which are not needed for the carotenoid biosynthesis. The plasmid p602-CAR has approx. 6.7 kb of genomic *Flavobacterium R1534* DNA containing besides all five carotenoid genes (approx. 4.9 kb), additional genomic DNA of 1.2 kb, located upstream of the crtZ translation start site and further 200 bp, located upstream of crtE transcription start. The crtZ, crtY, crtI and crtB genes were cloned downstream of the $P_{N25/0}$ promoter, a regulatable *E. coli* bacteriophage T5 promoter derivative, fused to a lac operator element, which is functional in *B. subtilis* [LeGrice, S.F.J., s.a.]. It is obvious that in the p602CAR construct, the distance of over 1200 bp between the $P_{N25/0}$ promoter and the transcription start site of crtZ is not optimal and will be improved at a later stage. An outline of the p602CAR construction is shown in figure 17. To ensure transcription of the crtE gene in *B. subtilis*, the vegI promoter [Moran et al., Mol. Gen. Genet. 186, 339-346

(1982); LeGrice et al., Mol. Gen. Genet. 204, 229-236 (1986)] was introduced upstream of this gene, resulting in the plasmid construct p602-CARVEG-E. The vegI promoter, which originates from sitel of the veg promoter complex described by [LeGrice et al., s.a.] has been shown to be functional in *E. coli* [Moran et al., s.a.]. To obtain this new construct, the plasmid p602CAR was digested with SalI and HindIII, and the fragment containing the complete crtE gene and most of the crtB coding sequence, was subcloned into the XhoI and HindIII sites of plasmid p205. The resulting plasmid p205CAR contains the crtE gene just downstream of the PvegI promoter. To reconstitute the carotenoid gene cluster of *Flavobacterium sp.* The following three pieces were isolated: PmeI/HindIII fragment of p205CAR, the HincII/XbaI fragment and the EcoRI/HindIII fragment of p602CAR and ligated into the EcoRI and XbaI sites of pBluescriptIIKS(+), resulting in the construct pBIIKS(+)-CARVEG-E. Isolation of the EcoRI-XbaI fragment of this latter plasmid and ligation into the EcoRI and XbaI sites of p602/22 gives a plasmid similar to p602CAR but having the crtE gene driven by the PvegI promoter. All the construction steps to get the plasmid p602CARVEG-E are outlined in figure 18. *E. coli* TG1 cells transformed with this plasmid synthesized zeaxanthin. In contrast *B. subtilis* strain 1012 transformed with the same constructs did not produce any carotenoids. Analysis of several zeaxanthin negative *B. subtilis* transformants always revealed, that the transformed plasmids had undergone severe deletions. This instability could be due to the large size of the constructs.

In order to obtain a stable construct in *B. subtilis*, the carotenoid genes were cloned into the Gram (+)/(-) shuttle vector pHP13 constructed by [Haima et al., s.a.]. The stability problems were thought to be omitted by 1) reducing the size of the cloned insert which carries the carotenoid genes and 2) reversing the orientation of the crtE gene and thus only requiring one promoter for the expression of all five genes, instead of two, like in the previous constructs. Furthermore, the ability of cells transformed by such a plasmid carrying the synthetic *Flavobacterium* carotenoid operon (SFCO), to produce carotenoids, would answer the question if a modular approach is feasible. Figure 19 summarizes all the construction steps and intermediate plasmids made to get the final construct pHP13-2PNZYIB-EINV. Briefly: To facilitate the following constructions, a vector pHP13-2 was made, by introducing a synthetic linker obtained with primer CS1 (SEQ ID NO: 19) and CS2 (SEQ ID NO: 20), between the HindIII and EcoRI sites of the shuttle vector pHP13. The intermediate construct pHP13-2CARVEG-E was constructed by subcloning the AflII-XbaI fragment of

p602CARVEG-E into the AflII and XbaI sites of pHP13-2. The next step consisted in the inversion of crtE gene, by removing XbaI and AvrII fragment containing the original crtE gene and replacing it with the XbaI-AvrII fragment of plasmid pBIIKS(+)-PCRRBScrtE. The resulting plasmid was named pHP13-2CARZYIB-EINV and represented the first construction with a functional SFCO. The intermediate construct pBIIKS(+)-PCRRBScrtE mentioned above, was obtained by digesting the PCR product generated with primers #100 (SEQ ID NO: 7) and #101 (SEQ ID NO: 8) with SpeI and SmaI and ligating into the SpeI and SmaI sites of pBluescriptIIKS(+). In order to get the crtZ transcription start close to the promoter $P_{N25/0}$ a triple ligation was done with the BamHI-SalI fragment of pHP13-2CARZYIB-EINV (contains four of the five carotenoid genes), the BamHI-EcoRI fragment of the same plasmid containing the $P_{N25/0}$ promoter and the EcoRI-SalI fragment of pBIIKS(+)-PCRRBScrtZ, having most of the crtZ gene preceded by a synthetic RBS. The aforementioned plasmid pBIISK(+)-PCRRBScrtZ was obtained by digesting the PCR product amplified with primers #104 (SEQ ID NO: 9) and #105 (SEQ ID NO: 10) with EcoRI and SalI and ligating into the EcoRI and SalI sites of pBluescriptIISK(+). In the resulting vector pHP13-2PN25ZYIB-EINV, the SFCO is driven by the bacteriophage T5 promoter $P_{N25/0}$, which should be constitutively expressed, due to the absence of a functional lac repressor in the construct [Peschke and Beuk, J. Mol. Biol. 186, 547-555 (1985)]. *E. coli* TG1 cells transformed with this construct produced zeaxanthin. Nevertheless, when this plasmid was transformed into *B. subtilis*, no carotenoid production could be detected. Analysis of the plasmids of these transformants showed severe deletions, pointing towards instability problems, similar to the observations made with the aforementioned plasmids.

Examples 6

Chromosome Integration Constructs

Due to the instability observed with the previous constructs we decided to integrate the carotenoid biosynthesis genes of *Flavobacterium sp.* into the genome of *B. subtilis* using the integration/expression vector pXI12. This vector allows the constitutive expression of whole operons after integration into the levan-sucrase gene (sacB) of the *B. subtilis* genome. The constitutive expression is driven by the vegI promoter and results in medium level expression. The plasmid pXI12-ZYIB-EINV4 containing the synthetic *Flavobacterium* carotenoid operon (SFCO) was constructed as follows: the NdeI-HincII fragment of pBIISK(+)-PCRRBScrtZ was

cloned into the NdeI and SmaI sites of pXI12 and the resulting plasmid was named pXI12-PCRcrtZ. In the next step, the BstEII-PmeI fragment of pHP13-2PN25ZYIB-EINV was ligated to the BstEII-PmeI fragment of pXI12-PCRcrtZ (see figure 20). *B. subtilis* transformed with the resulting construct pXI12-ZYIB-EINV4 can integrate the CAR genes either via a Campbell type reaction or via a reciprocal recombination. One transformant, BS1012::ZYIB-EINV4, having a reciprocal recombination of the carotenoid biosynthesis genes into the levan-sucrase gene was further analyzed (figure 21). Although this strain did not synthesize carotenoids, RNA analysis by Northern blots showed the presence of specific polycistronic mRNA's of 5.4 kb and 4.2 kb when hybridized to probe A (see figure 21, panel B). Whereas the larger mRNA has the expected message size, the origin of the shorter mRNA was unclear. Hybridization of the same Northern blot to probe B only detected the large mRNA fragment, pointing towards a premature termination of the transcription at the end of the crtB gene. The presence of a termination signal at this location would make sense, since in the original operon organisation in the *Flavobacterium sp. R1534* genome, the crtE and the crtB genes are facing each other. With this constellation a transcription termination signal at the 5' end of crtB would make sense, in order to avoid the synthesis of anti-sense RNA which could interfere with the mRNA transcript of the crtE gene. Since this region has been changed considerably with respect to the wild type situation, the sequences constituting this terminator may also have been altered resulting in a "leaky" terminator. Western blot analysis using antisera against the different crt enzymes of the carotenoid pathway, pointed towards the possibility that the ribosomal binding sites might be responsible for the lack of carotenoid synthesis. Out of the five genes introduced only the product of crtZ, the β-carotene hydroxylase was detectable. This is the only gene preceded by a RBS site, originating from the pXI12 vector, known to be functional in *B. subtilis*. Base pairing interactions between a mRNA's Shine-Dalgarno sequence [Shine and Delagarno, s. a.] and the 16S rRNA, which permits the ribosome to select the proper initiation site, have been proposed by [McLaughlin et al., J. Biol. Chem. 256,11283-11291 (1981)] to be much more stable in Gram-positive organisms (*B. subtilis*) than in Gram-negative organisms (*E. coli*). In order to obtain highly stable complexes we exchanged the RBS sites of the Gram-negative *Flavobacterium sp.*, preceding each of the genes crtY, crtI, crtB and crtE, with synthetic RBS's which were designed complementary to the 3' end of the *B. subtilis* 16S rRNA (see table 2). This exchange should allow an effective translation initiation of the different carotenoid genes in *B. subtilis*. The strategy chosen to construct this pXI12-ZYIB-

EINV4MUTRBS2C, containing all four altered sites is summarized in figure 20. In order to facilitate the further cloning steps in pBluescriptIIKS(+), additional restriction sites were introduced using the linker obtained with primer MUT7 and MUT8, cloned between the SalI and HindIII sites of said vector. The new resulting construct pBIIKS(+)-LINKER78 had the following restriction sites introduced: AvrII, PmlI, MulI, MunI, BamHI and SphI. The general approach chosen to create the synthetic RBS's upstream of the different carotenoid genes, was done using a combination of PCR based mutagenesis, where the genes were reconstructed using defined primers carrying the modified RBS sites, or using synthetic linkers having such sequences. Reconstitution of the RBS preceding the crtI and crtB genes was done by amplifying the crtI gene with the primers MUT2 (SEQ ID NO: 12) and MUT6 (SEQ ID NO: 15), which include the appropriate altered RBS sites. The PCR-I fragment obtained was digested with MunI and BamHI and ligated into the MunI and BamHI sites of pBIIKS(+)-LINKER78. The resulting intermediate construct was named pBIIKS(+)-LINKER78PCRI. Reconstitution of the RBS preceding the crtB gene was done using a small PCR fragment obtained with primer MUT3 (SEQ ID NO: 13), carrying the altered RBS site upstream of crtB, and primer CAR17 (SEQ ID NO: 16). The amplified PCR-F fragment was digested with BamHI and HindIII and sub cloned into the BamHI and HindIII sites of pBIIKS(+)-LINKER78, resulting in the construct pBIIKS(+)-LINKER78PCRF. The PCR-I fragment was cut out of pBIIKS(+)-LINKER78PCRI with BamHI and SapI and ligated into the BamHI and SapI sites of pBIIKS(+)-LINKER78PCRF. The resulting plasmid pBIIKS(+)-LINKER78PCRFI has the PCR-I fragment fused to the PCR-F fragment. This construct was cut with SalI and PmlI and a synthetic linker obtained by annealing of primer MUT9 (SEQ ID NO: 23) and MUT10 (SEQ ID NO: 24) was introduced. This latter step was done to facilitate the upcoming replacement of the original *Flavobacterium* RBS in the above mentioned construct. The resulting plasmid was named pBIIKS(+)-LINKER78PCRFIA. Assembling of the synthetic RBS's preceding the crtY and crtI genes was done by PCR, using primers MUT1 (SEQ ID NO: 11) and MUT5 (SEQ ID NO: 14). The amplified fragment PCR-G was made blunt end before cloning into the SmaI site of pUC18, resulting in construct pUC18-PCR-G. The next step was the cloning of the PCR-G fragment between the PCR-A and PCR-I fragments. For this purpose the PCR-G was isolated from pUC18-PCR-G by digesting with MunI and PmlI and ligated into the MunI and PmlI sites of pBIIKS(+)-LINKER78PCRFIA. This construct contains all four fragments, PCR-F, PCR-I, PCR-G and PCR-A, assembled adjacent to each other and containing three of the four artificial RBS sites (crtY, crtI and crtB). The exchange of the *Flavobacterium*

RBS's preceding the genes crtY, crtI and crtB by synthetic ones, was done by replacing the HindIII-SalI fragment of plasmid pXI12-ZYIB-EINV4 with the HindIII-SalI fragment of plasmid pBIIKS(+)-LINKER78PCRFIGA. The resulting plasmid pXI12-ZYIB-EINV4 MUTRBSC was subsequently transformed into *E. coli* TG1 cells and *B. subtilis 1012*. The production of zeaxanthin by these cells confirmed that the PCR amplified genes where functional. The *B. subtilis* strain obtained was named BS1012::SFCO1. The last *Flavobacterium* RBS to be exchanged was the one preceding the crtE gene. This was done using a linker obtained using primer MUT11 (SEQ ID NO: 25) and MUT12 (SEQ ID NO: 26). The wild type RBS was removed from pXI12-ZYIB-EINV4MUTRBS with NdeI and SpeI and the above mentioned linker was inserted. In the construct pXI12-ZYIB-EINV4MUTRBS2C all *Flavobacterium* RBS's have been replaced by synthetic RBS's of the consensus sequence AAAGGAGG- 7-8 N -ATG (see table 2). *E. coli* TG1 cells transformed with this construct showed that also this last RBS replacement had not interferred

Table 2

mRNA	nucleotide sequence
crtZ (SEQ ID NO: 48)	**AAAGGAGG**GUUUCAUAUGAGC
crtY (SEQ ID NO: 49)	**AAAGGAGG**ACACGUGAUGAGC
crtI (SEQ ID NO: 50)	**AAAGGAGG**CAAUUGAGAUGAGU
crtB (SEQ ID NO: 51)	**AAAGGAGG**AUCCAAUCAUGACC
crtE (SEQ ID NO: 52)	**AAAGGAGG**GUUUCUUAUGACG
B. subtilis	16S rRNA (SEQ ID NO: 53) 3'-**UCUUUCCUCCACUAG**
E. coli	16S rRNA (SEQ ID NO: 54) 3'- **AUUCCUCCACUAG**

Table 2: Nucleotide sequences of the synthetic ribosome binding sites in the constructs pXI12-ZYIB-EINV4MUTRBS2C, pXI12-ZYIB-EINV4MUTRBS2CCAT and pXI12-ZYIB-EINV4 MUTRBS2CNEO. Nucleotides of the Shine-Dalgarno sequence preceding the individual carotenoid genes which are complementary to the 3' ends of the 16S rRNA of *B. subtilis* are shown in bold. The 3' ends of the 16S

rRNA of *E. coli* is also shown as comparison. The underlined AUG is the translation start site of the mentioned gene.

with the ability to produce zeaxanthin. All the regions containing the newly introduced synthetic RBS's were confirmed by sequencing. *B. subtilis* cells were transformed with plasmid pXI12-ZYIB-EINV4MUTRBS2 and one transformant having integrated the SFCO by reciprocal recombination, into the levan-sucrase gene of the chromosome, was selected. This strain was named BS1012::SFCO2. Analysis of the carotenoid production of this strain show that the amounts zeaxanthin produced is approx. 40% of the zeaxanthin produced by *E. coli* cells transformed with the plasmid used to get the *B. subtilis* transformant. Similar was the observation when comparing the BS1012::SFCO1 strain with its *E. coli* counter part (30%). Although the *E. coli* cells have 18 times more carotenoid genes, the carotenoid production is only a factor of 2-3 times higher. More drastic was the difference observed in the carotenoid contents, between *E. coli* cells carrying the pZea4 construct in about 200 copies and the *E. coli* carrying the plasmid pXI12-ZYIB-EINV4MUTRBS2C in 18 copies. The first transformant produced 48x more zeaxanthin than the latter one. This difference seen can not only be attributed to the roughly 11 times more carotenoid biosynthesis genes present in these transformants. Contributing to this difference is probably also the suboptimal performance of the newly constructed SFCO, in which the overlapping genes of the wild type *Flavobacterium* operon were separated to introduce the synthetic RBS's. This could have resulted in a lower translation efficiency of the rebuild synthetic operon (e.g. due to elimination of putative translational coupling effects, present in the wild type operon).

In order to increase the carotenoid production, two new constructs were made, pXI12-ZYIB-EINV4MUTRBS2CNEO and pXI12-ZYIB-EINV4 MUTRBS2CCAT, which after the integration of the SFCO into the levan-sucrase site of the chromosome, generate strains with an amplifiable structure as described by [Janniere et al., Gene 40, 47-55 (1985)]. Plasmid pXI12-ZYIB-EINV4MUTRBS2CNEO has been deposited on May 25, 1995 at the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Germany) under accession No. DSM 10013. Such amplifiable structures, when linked to a resistance marker (e.g chloramphenicol, neomycin, tetracycline), can be amplified to 20-50 copies per chromosome. The amplifiable structure

consist of the SFCO, the resistance gene and the pXI12 sequence, flanked by direct repeats of the sac-B 3' gene (see figure 22). New strains having elevated numbers of the SFCO could now be obtained by selecting for transformants with increased level of resistance to the antibiotic. To construct plasmid pXI12-ZYIB-EINV4MUTRBS2CNEO, the neomycin resistance gene was isolated from plasmid pBEST501 with PstI and SmaI and subcloned into the PstI and EcoO1091 sites of the pUC18 vector. The resulting construct was named pUC18-Neo. To get the final construct, the PmeI - AatII fragment of plasmid pXI12-ZYIB-EINV4MUTRBS2C was replaced with the SmaI-AatII fragment of pUC18-Neo, containing the neomycin resistance gene. Plasmid pXI12-ZYIB-EINV4MUTRBS2CCAT was obtained as follows: the chloramphenicol resistance gene of pC194 was isolated by PCR using the primer pair cat3 (SEQ ID NO: 17) and cat4 (SEQ ID NO: 18). The fragment was digested with EcoRI and AatII and subcloned into the EcoRI and AatII sites of pUC18. The resulting plasmid was named pUC18-CAT. The final vector was obtained by replacing the PmeI-AatII fragment of pXI12-ZYIB-EINV4MUTRBS2C with the EcoRI-AatII fragment of pUC18-CAT, carrying the chloramphenicol resistance gene. Figure 23 summarizes the different steps to obtain aforementioned constructs. Both plasmids were transformed into *B. subtilis* strain 1012, and transformants resulting from a Campbell-type integration were selected. Two strains BS1012::SFCONEO1 and BS1012::SFCOCAT1 were chosen for further amplification. Individual colonies of both strains were independently amplified by growing them in different concentrations of antibiotics as described in the methods section. For the cat gene carrying strain, the chloramphenicol concentrations were 60, 80, 120 and 150 mg/ml. For the neo gene carrying strain, the neomycin concentrations were 160 and 180 mg/ml. In both strains only strains with minor amplifications of the SFCO's were obtained. In daughter strains generated from strain BS1012::SFCONEO1, the resistance to higher neomycin concentrations correlated with the increase in the number of SFCO's in the chromosome and with higher levels of carotenoids produced by these cells. A different result was obtained with daughter strains obtained from strain BS1012::SFCOCAT1. In these strains an increase up to 150 mg chloramphenicol/ml resulted, as expected, in a higher number of SFCO copies in the chromosome.

Example 7

Construction of CrtW containing plasmids and use for carotenoid production

Polymerase chain reaction based gene synthesis. The nucleotide sequence of the artificial crtW gene, encoding the β-carotene β-4-oxygenase of *Alcaligenes* strain PC-1, was obtained by back translating the amino acid sequence outlined in [Misawa, 1995], using the BackTranslate program of the GCG Wisconsin Sequence Analysis Package, Version 8.0 (Genetics Computer Group, Madison, WI, USA) and a codon frequency reference table of *E. coli* (supplied by the Bach Translate Program). The synthetic gene consisting of 726 nucleotides was constructed basically according to the method described by [Ye, 1992]. The sequence of the 12 oligonucleotides (crtW1 - crtW12) required for the synthesis are shown in Figure 25 (SEQ ID NO: 28). Briefly, the long oligonucleotides were designed to have short overlaps of 15-20 bases, serving as primers for the extension of the oligonucleotides. After four cycles a few copies of the full length gene should be present which is then amplified by the two terminal oligonucleotides crtW15 (SEQ ID NO: 55) and crtW26. The sequences for these two short oligonucleotides are for the forward primer crtW15 (5'-TATATCTAGAcatatgTCCGGTCGTAAA CCGG-3') and for the reverse primer crtW26 (SEQ ID NO: 56) (5'-TATAgaattccacgtgTCA AGCACGA CCACCGGTTTTAC G-3'), where the sequences matching the DNA templates are underlined. Small cap letters show the introduced restriction sites (*Nde*I for the forward primer and *Eco*RI and *Pml*I for the reverse primer) for the latter cloning into the pALTER-Ex2 expression vector.

Polymerase chain reaction. All twelve long oligonucleotides (crtW1-crtW12; 7 nM each) and both terminal primers (crtW15 and crtW26; 0.1 mM each) were mixed and added to a PCR reaction mix containing Expand™ High Fidelity polymerase (Boehringer, Mannheim) (3.5 units) and dNTP's (100 mM each). The PCR reaction was run for 30 cycles with the following profile: 94 °C for 1 min, 50 °C for 2 min and 72 °C for 3 min. The PCR reaction was separated on a 1% agarose gel, and the band of approx. 700 bp was excised and purified using the glass beads method (Geneclean Kit, Bio101, Vista, CA, USA). The fragment was subsequently cloned into the *Sma*I site of plasmid pUC18, using the Sure-Clone Kit (Pharmacia, Uppsala, Sweden). The sequence of the resulting crtW synthetic gene was verified by sequencing with the Sequenase Kit Version 1.0 (United States Biochemical, Cleveland, OH, USA). The crtW gene constructed by this method was found to contain minor errors, which were subsequently corrected by site-directed

mutagenesis.

Construction of plasmids. Plasmid pBIIKS(+)-CARVEG-E (see also Example 5) (Figure 26) contains the carotenoid biosynthesis genes (crtE, crtB, crtY, crtI and crtZ) of the Gram (-) bacterium *Flavobacterium* sp. strain R1534 WT (ATCC 21588) [Pasamontes, 1995 #732] cloned into a modified pBluescript II KS(+) vector (Stratagene, La Jolla, USA) carrying site I of the *B. subtilis* veg promoter [LeGrice, 1986 #806]. This constitutive promoter has been shown to be functional in *E. coli*. Transformants of *E. coli* strain TG1 carrying plasmid pBIIKS(+)-CARVEG-E synthesise zeaxanthin. Plasmid pALTER-Ex2-crtW was constructed by cloning the *Nde*I - *Eco*RI restricted fragment of the synthetic crtW gene into the corresponding sites of plasmid pALTER-Ex2 (Promega, Madison, WI). Plasmid pALTER-Ex2 is a low copy plasmid with the p15a origin of replication, which allows it to be maintained with ColE1 vectors in the same host. Plasmid pBIIKS-crtEBIYZW (Figure 26) was obtained by cloning the *Hind*III-*Pml*I fragment of pALTER-Ex2-crtW into the *Hind*III and the blunt end made *Mlu*I site obtained by a fill in reaction with Klenow enzyme, as described elsewhere in [Sambrook, 1989 #505]. Inactivation of the crtZ gene was done by deleting a 285 bp *Nsi*I-*Nsi*I fragment, followed by a fill in reaction and religation, resulting in plasmid pBIIKS-crtEBIY[DZ]W. Plasmid pBIIKS-crtEBIY[DZW] carrying the non-functional genes crtW and crtZ, was constructed by digesting the plasmid pBIIKS-crtEBIY[DZ]W with *Nde*I and *Hpa*I, and subsequent self religation of the plasmid after filling in the sites with Klenow enzyme. *E. coli* transformed with this plasmid had a yellow-orange colour due to the accumulation of β-carotene. Plasmid pBIIKS-crtEBIYZ[DW] has a truncated crtW gene obtained by deleting the *Nde*I - *Hpa*I fragment in plasmid pBIIKS-crtEBIYZW as outlined above. Plasmids pALTER-Ex2-crtEBIY[DZW] and pALTER-Ex2-crtEBIYZ[DW], were obtained by isolating the *Bam*HI-*Xba*I fragment from pBIIKS-crtEBIY[DZW] and pBIIKS-crtEBIYZ[DW], respectively and cloning them into the *Bam*HI and *Xba*I sites of pALTER-Ex2. The plasmid pBIIKS-crtW was constructed by digesting pBIIKS-crtEBIYZW with *Nsi*I and *Sac*I, and self-religating the plasmid after recessing the DNA overhangs with Klenow enzyme. Figure 27 compiles the relevant inserts of all the plasmids used in this paper.

Carotenoid analysis. E. coli TG-1 transformants carrying the different plasmid constructs were grown for 20 hours in Luria-Broth medium supplemented with antibiotics (ampicillin 100 mg/ml, tetracyclin 12.5 mg/ml) in shake flasks at 37°C

and 220 rpm. Carotenoids were extracted from the cells with acetone. The acetone was removed in vacuo and the residue was re dissolved in toluene. The coloured solutions were subjected to high-performance liquid chromatography (HPLC) analysis which was performed on a Hewlett-Packard series 1050 instrument. The carotenoids were separated on a silica column Nucleosil Si - 100, 200 x 4 mm, 3m. The solvent system included two solvents: hexane (A) and hexane/THF , 1:1 (B). A linear gradient was applied running from 13 to 50 % (B) within 15 minutes. The flow rate was 1.5 ml/min. Peaks were detected at 450 nm by a photo diode array detector. The individual carotenoid pigments were identified by their absorption spectra and typical retention times as compared to reference samples of chemically pure carotenoids, prepared by chemical synthesis and characterised by NMR, MS and UV-Spectra. HPLC analysis of the pigments isolated from *E. coli* cells transformed with plasmid pBIIKS-crtEBIYZW, carrying besides the carotenoid biosynthesis genes of *Flavobacterium* sp. strain R1534, also the crtW gene encoding the β-carotene ketolase of *Alcaligenes* PC-1 [Misawa, 1995 #670] gave the following major peaks identified as: b-cryptoxanthin, astaxanthin, adonixanthin and zeaxanthin, based on the retention times and on the comparison of the absorbance spectra to given reference samples of chemically pure carotenoids. The relative amount (area percent) of the accumulated pigment in the E. coli transformant carrying pBIIKS-crtEBIYZW is shown in Table 3 ["CRX": cryptoxanthin; "ASX": astaxanthin; "ADX": adonixanthin; "ZXN": zeaxanthin; "ECM": echinenone; "MECH": 3-hydroxyechinenone, "CXN": canthaxanthin]. The Σ of the peak areas of all identified carotenoids was defined as 100%. Numbers shown in Table 3 represent the average value of four independent cultures for each transformant. In contrast to the aforementioned results, *E. coli* transformants carrying the same genes but on two plasmids namely, pBIIKS-crtEBIYZ[DW] and pALTER-Ex2-crtW, showed a drastical drop in adonixanthin and a complete lack of astaxanthin pigments (Table 3), whereas the relative amount of zeaxanthin (%) had increased. Echinenone, hydroxyechinenone and canthaxanthin levels remained unchanged compared to the transformants carrying all the crt genes on the same plasmid (pBIIKS-crtEBIYZDW). Plasmid pBIIKS-crtEBIYZ[DW] is a high copy plasmid carrying the functional genes of crtE, crtB, crtY, crtI, crtZ of *Flavobacterium* sp. strain R1534 and a truncated, non-functional version of the crtW gene, whereas the functional copy of the crtW gene is located on the low copy plasmid pALTER-Ex2-crtW. To analyze the effect of overexpression of the crtW gene with respect to the crtZ gene, *E. coli* cells were co-transformed with plasmid pBIIKS-crtW carrying the crtW gene on the high copy plasmid pBIIKS-

crtW and the low copy construct pALTER-Ex2-crtEBIYZ[DW], encoding the *Flavobacterium* crt genes. Pigment analysis of these transformants by HPLC monitored the presence of β-carotene, cryptoxanthin, astaxanthin, adonixanthin, zeaxanthin, 3-hydroxyechine-none and minute traces of echinenone and canthaxanthin (Table 3).

Transformants harbouring the crtW gene on the low copy plasmid pALTER-Ex2-crtW and the genes crtE, crtB, crtY and crtI on the high copy plasmid pBIIKS-crtEBIY[DZW] expressed only minor amounts of canthaxanthin (*6 %*) but high levels of echinenone (*94%*), whereas cells carrying the crtW gene on the high copy plasmid pBIIKS-crtW and the other crt genes on the low copy construct pALTER-Ex2-crtEBIY[DZW], had *78.6 %* and *21.4 %* of echinenone and canthaxanthin, respectively (Table 3).

Table 3

plasmids	CRX	ASX	ADX	ZXN	ECH	HECH	CXN
pBIIKS-crtEBIYZW	1.1	2.0	44.2	52.4	<1	<1	<1
pBIIKS-crtEBIYZ[W]+ pALTER-Ex2-crtW	2.2	-	25.4	72.4	<1	<1	<1
pBIIKS-crtEBIY[Z]W	-	-	-	-	66.5	-	33.5
pBIIKS-crtEBIY[ZW]+ pBIIKS-crtW	-	-	-	-	94	-	6

Example 8

Selective carotenoid production by using the crtW and crtZ genes of the Gram negative bacterium E-396.

In this section we describe *E. coli* transformants which accumulate only one (canthaxanthin) or two main carotenoids (astaxanthin, adonixanthin) and minor amounts of adonirubin, rather than the complex variety of carotenoids seen in most carotenoid producing bacteria [Yokoyama et al., Biosci. Biotechnol. Biochem. 58:1842-1844 (1994)] and some of the *E. coli* transformants shown in Table 3. The

ability to construct strains producing only one carotenoid is a major step towards a successful biotechnological carotenoid production process. This increase in the accumulation of individual carotenoids accompanied by a decrease of the intermediates, was obtained by replacing the crtZ of *Flavobacterium* R1534 and/or the synthetic crtW gene (see example 5) by their homologous genes originating from the astaxanthin producing Gram negative bacterium E-396 (FERM BP-4283) [Tsubokura et al., EP-application 0 635 576 A1]. Both genes, $crtW_{E396}$ and $crtZ_{E396}$, were isolated and used to construct new plasmids as outlined below.

Isolation of a putative fragment of the crtW gene of strain E-396 by the polymerase chain reaction. Based on protein sequence comparison of the crtW enzymes *of Agrobacterium aurantiacum, Alcaligenes PC-1* (WO95/18220) [Misawa et al., J.Bacteriol. 177: 6575-6584 (1995)] and *Haematococcus pluvialis* [Kajiwara et al., Plant Mol. Biol. 29:343-352 (1995)][Lotan et al., FEBS letters, 364:125-128 (1995)], two regions named I and II, having high amino acid conservation and located approx. 140 amino acids appart, were identified and chosen to design the degenerate PCR primers shown below. The N-terminal peptide HDAMHG (region I) was used to design the two 17-mer degenerate primer sequences crtW100 (SEQ ID NO: 57) and crtW101 (SEQ ID NO: 58):

crtW100 (SEQ ID NO: 57): 5'-CA(C/T)GA(C/T)GC(A/C)ATGCA(C/T)GG-3'

crtW101 (SEQ ID NO: 58): 5'-CA(C/T)GA(C/T)GC(G/T)ATGCA(C/T)GG-3'

The C-terminal peptide H(W/H)EHH(R/L) corresponding to region II was used design the two 17-mer degenerate primer with the antisense sequences crtW105 (SEQ ID NO: 59) and crtW106 (SEQ ID NO: 60):

crtW105 (SEQ ID NO: 59): 5'-AG(G/A)TG(G/A)TG(T/C)TC(G/A)TG(G/A)TG-3'

crtW106 (SEQ ID NO: 60): 5'-AG(G/A)TG(G/A)TG(T/C)TCCCA(G/A)TG-3'

Polymerase chain reaction. PCR was performed using the GeneAmp Kit (Perkin Elmer Cetus) according to the manufacturer's instructions. The different PCR reactions contained combinations of the degenerate primers (crtW100/crtW105 or crtW100/crtW106 or crtW101/crtW105 or crtW101/crtW106) at a final concentration of 50 pM each, together with genomic DNA of the bacterium E-396 (200 ng) and 2.5 units of Taq polymerase. In total 35 cycles of PCR were performed with the following cycle profile: 95 °C for 30 sec, 55 °C for 30 sec, 72

°C for 30 sec. PCR reactions made with the following primer combinations crtW100/crtW105 and crtW101/crtW105 gave PCR amplification products of approx. 500 bp which were in accordance with the expected fragment size. The 500 bp fragment, JAPclone8, obtained in the PCR reaction using primers crtW101 (SEQ ID NO: 58) and crtW105 (SEQ ID NO: 59) was excised from an 1.5% agarose gel and purified using the GENECLEAN Kit and subsequently cloned into the *Sma*I site of pUC18 using the Sure-Clone Kit, according to the manufacturer's instructions. The resulting plasmid was named pUC18-JAPclone8 and the insert was sequenced. Comparison of the determined sequence to the crtW gene of *Agrobacterium aurantiacum* (GenBank accession n° D58420) published by Misawa *et al.* in 1995 (WO95/18220) showed 96% identity at the nucleotide sequence level, indicating that both organisms might be closely related.

Isolation of the crt cluster of the strain E- 396. Genomic DNA of E-396 was digested overnight with different combinations of restrictions enzymes and separated by agarose gel electrophoresis before transferring the resulting fragments by Southern blotting onto a nitrocellulose membrane. The blot was hybridised with a ^{32}P labelled 334 bp fragment obtained by digesting the aforementioned PCR fragment JAPclone8 with *Bss*HII and *Mlu*I. An approx. 9,4kb *Eco*RI/*Bam*HI fragment hybridizing to the probe was identified as the most appropiate for cloning since it is long enough to potentially carry the complete *crt* cluster. The fragment was isolated and cloned into the *Eco*RI and *BamH*I sites of pBluescriptIIKS resulting in plasmid pJAPCL544 (Fig. 29). Based on the sequence of the PCR fragment JAPclone8, two primers were synthesized to obtain more sequence information left and right hand of this fragment. Fig. 30 shows the sequence obtained containing the $crtW_{E396}$ (from nucleotide 40 to 768) and $crtZ_{E396}$ (SEQ ID NO: 33) (from nucleotide 765 to 1253) genes of the bacterium E-396. The nucleotide sequence of the $crtW_{E396}$ (SEQ ID NO: 30) gene is shown in Fig. 31 (SEQ ID NO: 31) and the encoded amino acid sequence in Fig. 32 (SEQ ID NO: 32). The nucleotide sequence of the $crtZ_{E396}$ gene is shown in Fig. 33 (SEQ ID NO: 33) and the corresponding amino acid sequence in Fig. 34 (SEQ ID NO: 34). Comparison to the $crtW_{E396}$ gene of E-396 to the crtW gene of *A. aurantiacum* showed 97 % identity at the nucleotide level and 99 % identity at the amino acid level. For the crtZ gene the values were 98 % and 99 %, respectively.

Construction of plasmids: Both genes, $crtW_{E396}$ and $crtZ_{E396}$, which are adjacent in the genome of E-396, were isolated by PCR using primer crtW107 and

crtW108 and the ExpandTM High Fidelity PCR system of Boehringer Mannheim, according to the manufacturer's recommendations. To facilitate the subsequent cloning steps (see section below) the primer crt107 (SEQ ID NO: 61) (5'-ATCATATGAGCGCACATGCCCTGCCCAAGGC-3') contains an artificial *Nde*I site (underlined sequence) spanning the ATG start codon of the $crtW_{E396}$ gene and the reverse primer crtW108 (SEQ ID NO: 62) (5'-ATCTCGAGTCACGTGCGC TCCTGCGCCTCGGCC-3') has an *Xho*I site (underlined sequence) just downstream of the TGA stop codon of the $crtZ_{E396}$ gene. The final PCR reaction mix had 10 pM of each primer, 2.5 mg genomic DNA of the bacterium E-396 and 3.5 units of the TaqDNA/Pwo DNA polymerase mix. In total 35 cycles were performed with the following cycle profile: 95 °C, 1 min; 60 °C, 1 min; 72 °C 1min 30 sec. The PCR product of approx. 1250bp was isolated from the 1% agarose gel and purified using GENECLEAN before ligation into the *Sma*I site pUC18 using the Sure-Clone Kit. The resulting construct was named pUC18-E396crtWZPCR (Fig. 35). The functionality of both genes was tested as follows. The $crtW_{E396}$ and $crtZ_{E396}$ gene were isolated from plasmid pUC18-E396crtWZPCR with *Nde*I and *Xho*I and cloned into the *Nde*I and *Sal*I site of plasmid pBIIKS-crtEBIYZW resulting in plasmid pBIIKS-crtEBIY[E396WZ] (Fig. 36). *E. coli* TG1 cells transformed with this plasmid produced astaxanthin, adonixanthin and adonirubin but no zeaxanthin (Table 4).

Plasmid pBIIKS-crtEBIY[E396W]DZ has a truncated non-functional crtZ gene. Fig. 37 outlines the construction of this plasmid. The PCR reaction was run as outlined elsewhere in the text using primers crtW113 (SEQ ID NO: 63)/crtW114 (SEQ ID NO: 64) and 200 ng of plasmid pUC18-JAPclone8 as template using 20 cycles with the following protocol: 95 °C, 45 sec/ 62 °C, 20 sec/ 72 °C, 20 sec)

primer crtW113 (SEQ ID NO: 63)
(5'-ATATACATATGGTGTCCCCCTTGGTGCGGGTGC-3')

primer crtW114 (SEQ ID NO: 64)
(5'-TATGGATCCGACGCGTTCCCGGACCGCCACAATGC-3')

The resulting 150 bp fragment was digested with *BamH*I and *Nde*I and cloned into the corresponding sites of pBIISK(+)-PCRRBScrtZ resulting in the construct pBIISK(+)-PCRRBScrtZ-2.The final plasmid carrying the genes crtE, crtB, crtI, crtY of *Flavobacterium*, the $crtW_{E396}$ gene of E-396 and a truncated non-functional crtZ gene of *Flavobacterium* was obtained by isolating the *Mlu*I/*Nru*I fragment (280 bp) of pBIISK(+)-PCRRBScrtZ-2 and cloning it, into the *Mlu*I/*Pml*I sites of

plasmid pBIIKS-crtEBIY[E396WZ]. *E. coli* cells transformed with this plasmid produced 100% canthaxanthin (Table 4; "CRX": cryptoxanthin; "ASX": astaxanthin; "ADX": adonixanthin; "ZXN": zeaxanthin; "ECH": echinenone; "HECH": 3-hydroxyechinenone; "CXN": canthaxanthin; "BCA": β-carotene; "ADR": adonirubin; Numbers indicate the % of the individual carotenoid of the total carotenoids produced in the cell.).

Table 4

plasmid	CRX	ASX	ADX	ZXN	ECH	HECH	CXN	BCA	ADR
pBIIKScrtEBIYZW	1.1	2.0	44.2	52.4	<1	<1	<1		
pBIIKS-crtEBIY[E396WZ]		74.4	19.8						5.8
pBIIKS-crtEBIY[E396W]DZ							100		

The results of *E. coli* transformants carrying pBIIKScrtEBIYZW (see example 7) are also shown in Table 4 to indicate the dramatic effect of the new genes $crtW_{E396}$ and $crtZ_{E396}$ on the carotenoids produced in these new transformants.

Example 9

Cloning of the remaining *crt* genes of the Gram negative bacterium E-396.

TG1 *E. coli* transformants carrying the pJAPCL544 plasmid did not produce detectable quantities of carotenoids (results not shown). Sequence analysis and comparison of the 3' (*BamHI* site) of the insert of plasmid pJAPCL544, to the *crt* cluster of *Flavobacterium* R1534 showed that only part of the C-terminus of the crtE gene was present. This result explained the lack of carotenoid production in the aforementioned transformants. To isolate the missing N-terminal part of the gene, genomic DNA of E-396 was digested by 6 restrictions enzymes in different combinations : *Eco*RI, *Bam*HI, *Pst*I, *Sac*I, *Sph*I and *Xba*I and transferred by the Southern blot technique to nitrocellulose. Hybridization of this membrane with the ^{32}P radio-labelled probe (a 463 bp *PstI-BamHI* fragment originating from the 3'

end of the insert of pJAPCL544 (Fig. 29) highlighted a ~1300 bp-long *Pst*I-*Pst*I fragment. This fragment was isolated and cloned into the *Pst*I site of pBSIIKS(+) resulting in plasmid pBSIIKS-#1296. The sequence of the insert is shown in Fig. 38 (SEQ ID NO: 35) (small cap letters refer to new sequence obtained. Capital letters show the sequence also present in the 3' of the insert of plasmid pJAPCL544). The complete crtE gene has therefore a length of 882 bp (see Fig. 39) and encodes a GGPP synthase of 294 amino acids (Fig. 40) (SEQ ID NO: 37). The crtE enzyme has 38 % identity with the crtE amino acid sequence of *Erwinia herbicola* and 66 % with *Flavobacterium* R1534 WT.

Construction of plasmids. To have a plasmid carrying the complete *crt* cluster of E-396, the 4.7 kb *Mlu*I/*Bam*HI fragment encoding the genes crtW, crtZ, crtY, crtI and crtB was isolated from pJAPCL544 and cloned into the *Mlu*I/*Bam*HI sites of pUC18-E396crtWZPCR (see example 8). The new construct was named pE396CARcrtW-B (Fig. 41) and lacked the N-terminus of the crtE gene. The missing C-terminal part of the crtE gene was then introduced by ligation of the aforementioned *Pst*I fragment of pBIIKS-#1296 between the *Pst*I sites of pE396CARcrtW-B. The resulting plasmid was named pE396CARcrtW-E (Fig. 41). The carotenoid distribution of the *E. coli* transformants carrying aforementioned plasmid were: adonixanthin (65%), astaxanthin (8%) and zeaxanthin (3%). The % indicated reflects the proportion of the total amount of carotenoid produced in the cell.

Example 10

Astaxanthin and adonixanthin production *in Flavobacterium* R1534

Among bacteria *Flavobacterium* may represent the best source for the development of a fermentative production process for 3*R*, 3*R*' zeaxanthin. Derivatives of *Flavobacterium* sp. strain R1534, obtained by classical mutagenesis have attracted in the past two decades wide interest for the development of a large scale fermentative production of zeaxanthin, although with little success. Cloning of the carotenoid biosynthesis genes of this organism, as outlined in example 2, may allow replacement of the classical mutagenesis approach by a more rational one, using molecular tools to amplify the copy number of relevant genes, deregulate their expression and eliminate bottlenecks in the carotenoid biosynthesis pathway. Furthermore, the introduction of additional heterologous genes (e.g. crtW) will result in the production of carotenoids normally not

synthesised by this bacterium (astaxanthin, adonirubin, adonixanthin, canthaxanthin, echinenone). The construction of such recombinant *Flavobacterium* R1534 strains producing astaxanthin and adonixanthin will be outlined below.

Gene transfer into *Flavobacterium* sp.

Plasmid transfer by conjugative mobilization. For the conjugational crosses we constructed plasmid pRSF1010-Amp^r, a derivative of the small (8.9 kb) broad host range plasmid RSF1010 (IncQ incompatibility group) [Guerry et al., J. Bacteriol. 117:619-630 (1974)] and used E. coli S17-1 as the mobilizing strain [Priefer et al., J. Bacteriol. 163:324-330 (1985)]. In general any of the IncQ plasmids (e.g. RSF1010, R300B, R1162) may be mobilized into rifampicin resistant Flavobacterium *if the* transfer functions are provided by plasmids of the IncP1 group (e.g. R1, R751).

Rifampicin resistant (Rif^r) *Flavobacterium* R1534 cells were obtained by selection on 100 mg rifampicin/ml. One resistant colony was picked and a stock culture was made. The conjugation protocol was as follows:

Day 1:

- grow 3 ml culture of *Flavobacterium* R1534 Rif^r for 24 hours at 30 °C in Flavobacter medium (F-medium) (see example 1)

- grow 3 ml mobilizing *E. coli* strain carrying the mobilizable plasmid O/N at 37 °C in LB medium. (e.g *E. coli* S17-1 carrying pRSF1010-Amp^r or *E. coli* TG-1 cells carrying R751 and pRSF1010-Amp^r)

Day 2:

- pellet 1 ml of the *Flavobacterium* R1534 Rif^r cells and resuspend in 1ml of fresh F-medium.

- pellet 1 ml of *E. coli* cells (see above) and resuspend in 1 ml of LB medium.

-donor and recipient cells are then mixed in a ratio of 1:1 and 1: 10 in an Eppendorf tube and 30 ml are then applied onto a nitrocellulose filter plated on agar plates containing F-medium and incubated O/N at 30°C.

Day 3:

- the conjugational mixtures were washed off with F-medium and plated on F-medium containing 100 mg rifampicin and 100 mg ampicillin/ml for selection of transconjugants and inhibition of the donor cells.

Day 6-8:

- Arising clones are plated once more on F-medium containing 100 mgRif and 100 mg Amp/ml before analysis.

Plasmid transfer by electroporation. The protocol for the eletroporation is as follows:

1. add 10 ml of O/N culture of *Flavobacterium* sp. R1534 into 500 mlF-medium and incubate at 30°C until OD600=0.8-0.1
2. harvest cells by centrifugation at 4000g at 4°C for 10 min.
3. wash cells in equal volume of ice-cold deionized water (2 times)
4. resuspend bacterial pellet in 1 ml ice-cold deionized water
5. take 50 ml of cells for electroporation with 0.1 mg of plasmid DNA
6. electroporation was done using field strengths between 15 and 25 kV/cm and 1-3 ms.
7. after electroporation cells were immediately diluted in 1 ml of F-medium and incubated for 2 hours at 30°C at 180 rpm before plating on F-medium plates containing the respective selective antibioticum.

Plasmid constructions: Plasmid pRSF101-Amp^r was obtained by cloning the Amp^r gene of pBR322 between the *EcoRI/NotI* sites of RSF1010. The Amp^r gene originates from pBR322 and was isolated by PCR using primers AmpR1 (SEQ ID NO: 65) and AmpR2 (SEQ ID NO: 66) as shown in Fig. 42.

AmpR1 (SEQ ID NO: 65):

5'-TATATCGGCCGACTAGTAAGCTTCAAAAAGGATCTTCACCTAG-3' the underlined sequence contains the introduced restriction sites for *EagI, SpeI* and

*Hind*III to facilitate subsequent constructions.

AmpR2 (SEQ ID NO: 66):

5'-ATATGAATTCAATAATATTGAAAAAGGAAG-3' the underlined sequence corresponds to an introduced *Eco*RI restriction site to facilitate cloning into RSF1010 (see Fig. 42).

The PCR reaction mix had 10 pM of each primer (AmpR1 (SEQ ID NO: 65)/AmpR2 (SEQ ID NO: 66)), 0.5 mg plasmid pBR322 and 3.5 units of the TaqDNA/Pwo DNA polymerase mix . In total 35 amplification cycles were made with the profile: 95 °C, 45 sec; 59 °C, 45 sec, 72 °C, 1 min. The PCR product of approx. 950 was extracted once with phenol/chloroform and precipitated with 0.3 M NaAcetate and 2 vol. Ethanol. The pellet was resuspended in H_2O and digested with *Eco*RI and *Eag*I O/N. The digestion was separated by electrophoresis and the fragment isolated from the 1% agarose gel and purified using GENECLEAN before ligation into the *Eco*RI and *Not*I sites of RSF1010. The resulting plasmid was named pRSF1010-Amp^r (Fig. 42).

Plasmid RSF1010-Ampr-crt1 was obtained by isolating the *Hind*III/*Not*I fragment of pBIIKS-crtEBIY[E396WZ] and cloning it between the *Hind*III/EagI sites of RSF1010-Amp^r (Fig. 43). The resulting plasmid RSF1010-Ampr-crt1 carries $crtW_{E396}$, $crtZ_{E396}$, crtY genes and the N-terminus of the crtI gene (non-functional). Plasmid RSF1010-Ampr-crt2 carrying a complete crt cluster composed of the genes $crtW_{E396}$and $crtZ_{E396}$ of E-396 and the crtY, crtI, crtB and crtE of *Flavobacterium* R1534 was obtained by isolating the large *Hind*III/*Xba*I fragment of pBIIKS-crtEBIY[E396WZ] and cloning it into the *Spe*I/*Hind*III sites of RSF1010-Amp^r (Fig. 43).

Flavobacterium R1534 transformants carrying either plasmid RSF1010-Amp^r, Plasmid RSF1010-Amp^r-crt1 or Plasmid RSF1010-Amp^r-crt2 were obtained by conjugation as outlined above using *E. coli* S17-1 as mobilizing strain.

Comparison of the carotenoid production of two Flavobacterium transformants. Overnight cultures of the individual transformants were diluted into 20 ml fresh F-medium to have a final starting OD600 of 0.4. Cells were harvested after growing for 48 hours at 30 °C and carotenoid contents were analysed as outlined in example 7. Table 5 shows the result of the three control cultures *Flavobacterium* [R1534 WT], [R1534 WT RifR] (rifampicin resistant) and

[R1534WT Rifr RSF1010-AmpR] (carries the RSF1010-Amp^r plasmid) and the two transformants [R1534 WT RSF1010-AmpR-crt1] and [R1534 WT RSF1010-AmpR-crt2]. Both latter transformants are able to synthesise astaxanthin and adonixanthin but little zeaxanthin. Most interesting is the [R1534 WT RSF1010-AmpR-crt2] *Flavobacterium* transformant which produces approx. 4 times more carotenoids than the R1534 WT. This increase in total carotenoid production is most likely due to the increase of the number of carotenoid biosynthesis clusters present in these cell (e.g. corresponds to the total copy number of plasmids in the cell).

Table 5

Transformant	**carotenoids % of total dry weight**	**total carotenoid content in % of dry weight**
R1534 WT	0.039% b-Carotin 0.001% b-Cryptoxanthin 0.018% Zeaxanthin	0.06%
R1534 Rifr	0.036% b-Carotin 0.002% b-Cryptoxanthin 0.022% Zeaxanthin	0.06%
R1534 Rifr [RSF1010-Ampr]	0.021% b-Carotin 0.002% b-Cryptoxanthin 0.032% Zeaxanthin	0.065%
R1534 Rifr [RSF1010-Ampr-crt1]	0.022% Astaxanthin 0.075% Adonixanthin 0.004% Zeaxanthin	0.1%
R1534 Rifr [RSF1010-Ampr-crt2]	0.132% b-Carotin 0.006% Echinenon 0.004% Hydroxyechinenon 0.003% b-Cryptoxanthin 0.044% Astaxanthin 0.039% Adonixanthin 0.007% Zeaxanthin	0.235%